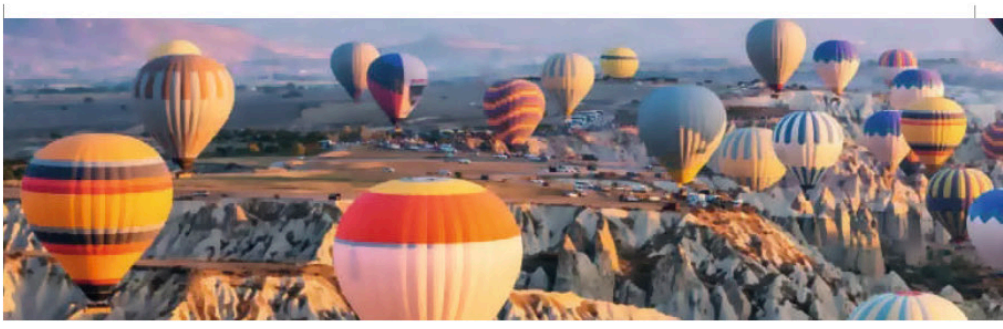


ESERA 2023 Cappadocia

The 15th ESERA

EUROPEAN SCIENCE EDUCATION RESEARCH ASSOCIATION

Proceedings Book Series-I

Editors
Metin Sardag
Gokhan Kaya
Mehmet Sogut

nobel yilmsel

EUROPEAN SCIENCE EDUCATION RESEARCH ASSOCIATION

ESERA

The 15th Conference of the European Science Education Research Association (ESERA)

ESERA 2023 Cappadocia

Proceedings Book Series-I

Strands 1, 2, 4, 14, 17, 18, 19

ESERA 2023 Conference
Cappadocia, Türkiye
August 28- September 1, 2023

Editors:
Metin Sardag, Gokhan Kaya, Mehmet Sogut

Foreword

It is my pleasure to introduce the proceedings of the 15th Conference of the European Science Education Research Association (ESERA) in Cappadocia, Türkiye. The conference was the first physical conference following the Covid-19 pandemic, and as such it was a wonderful opportunity for the ESERA community to reconnect in person and to learn about our respective research across Europe and beyond. I am particularly delighted that ESERA was able to provide a platform for early career researchers and new members who joined the ESERA family during this conference. Based on my interactions with the community at the conference, the event was intellectually stimulating and socially pleasant experience for all.

The planning of the conference began when we were still in the midst of the pandemic, slowly emerging from a very challenging period globally. The three universities involved in the organisation of the conference - Hacettepe University, Gazi University, and Nevşehir Hacı Bektaş Veli University - provided enormous support throughout but it is important to note particularly the period following the devastating earthquakes in February 2023. On behalf of the ESERA Board and the ESERA community, I thank the leadership of the Local Organising Committee, namely Gultekin Cakmakci, Mehmet Fatih Tasar, and Mustafa Hilmi Colakoglu for the effective management of the preparations under such extraordinary circumstances.

The theme of the conference was "Connecting Science Education with Cultural Heritage", one of the main goals of UNESCO, the United Nations Educational, Scientific and Cultural Organization. As a culturally diverse country with a rich heritage, Türkiye provided a brilliant context to consider research in science education in diverse social learning environments.

The five books of the conference proceedings have been edited by Gokhan Kaya, Metin Sardag, Mehmet Sogut, Tugba Ecevit, Kibar Gul, Orhan Curaoglu, Bahadir Altintas and Ismail Donmez. We acknowledge and thank them for their editorial input into the ESERA 2023 conference proceedings.

I hope that the conference proceedings will serve not only as a memory of a useful conference but also a forum where the newly established professional networks can flourish further, and help enhance our understanding of improving science education through research.

Sibel Erduran

ESERA President (2019-2023)

University of Oxford, UK

Foreword

A very warm welcome to the European Science Education Research Association (ESERA). The preparation for the conference began more than 2 years ago. We received around 1300 submissions. After a review process, we had 695 Papers, 54 Symposia, 2 Panels, 20 Workshops, 150 Posters with 3 plenary speakers and 2 ESERA community plenary speakers. With on-site registration, there were around 1100 participants. After a review process, 195 manuscripts were accepted to be part of these five eProceedings.

As always ESERA supports early-career researchers. We think that everything needs to be done to keep their enthusiasm fresh and alive for developing a science education culture. Drawing upon that we had five workshops for early-career researchers and practitioners as part of the programme for the ESERA 2023 conference. That brings fresh enthusiasm to the new and young generation of researchers and practitioners in STEM education around the globe. Providing opportunities for early-career researchers to serve on the editorial board for the development of these five eProceedings is one of the outcomes of that initiative.

Spanning Europe and Asia, it can be said that Türkiye has been the meeting place of many peoples and cultures throughout the centuries. Having colleagues from over 65 countries at this conference in Cappadocia, a historical region in Central Anatolia, enabled us to show our way of mutual understanding, respect, and cooperation which is also much needed in today's world. Accordingly, the theme of the ESERA 2023 conference was "Connecting Science Education with Cultural Heritage".

The ESERA 2023 conference was a great success with all conference participants and science education around the world. These five eProceedings will allow us to further debate in the field. We look forward to these valuable interactions and experiences, developing partnerships, and significant outcomes of the conference.

On behalf of the Local Organizing Committee of the ESERA 2023 Conference, I would like to thank everyone who has contributed to this conference and eProceedings in different capacities.

Looking forward to seeing you all at the ESERA 2025 Conference in Copenhagen, Denmark!

Gultekin Cakmakci

ESERA 2023 Conference President
Hacettepe University STEM & Maker Lab, Türkiye

Connecting Science Education with Cultural Heritage

Proceedings of the ESERA 2023 Conference

Publication No.: 2276

ISBN: 978-625-393-909-0

Edition Number: 1st Edition, September 2024

Editors for Proceedings Book Series-I:

Metin Sardag, Gokhan Kaya, Mehmet Sogut

Co-editors for Proceedings Book Series-I:

Ana Sofia Afonso, Massimiliano Malgieri, Regina Soobard, Radu Bogdan Toma, María Napal, Leticia Garcia-Romano, Jens Dolin, Lukas Rokos, Bodil Sundberg, Christina Siry, Carla Morais, İbrahim Delen, Federico Corni, May Lee

The Proceedings of ESERA 2023 is an electronic publication for revised and extended papers presented at the ESERA 2023 conference in Cappodocia, Türkiye during the August 28- September 1, 2023. All papers in the eProceedings correspond to communications submittedand accepted for the ESERA 2023 conference. All proposals to the conference went through a blind review by two or three reviewers prior to being accepted to the conference. A total of 695 papers, 54 symposia, 2 panel, 20 workshops, 150 posters were presented at the conference and in total 195 papers are included in the eProceedings.

The authors were asked to produce updated versions of their papers and take into account the discussion that took place after the presentation and the suggestions received from other participants at the conference. On the whole, the eProceedings presents acomprehensive overview of ongoing studies in Science Education Research in Europe and beyond. This book represents the current interests and areas of emphasis in the ESERA community at the end of 2023.

The Proceedings book series divided five different books eighteen parts that represent papers presented across 20 strands at the ESERA 2023 conference. The stand chairs for ESERA 2023 co-edited the corresponding part for each strand 1 to 20 and. All formats of presentation (single oral, interactive poster, ICT demonstration/workshop and symposium) used during the conference were eligible to be submitted to the eProceedings.

The co-editors carried out a review of the updated versions of the papers that were submitted after the conference at the end of ESERA 2023. The editors and

co-editors do not necessarily endorse or share the ideas and views presented in or implied by the papers included in this book.

The appropriate APA style for referencing this *Proceedings Book Series-I* is as follows:

Sardag, M., Kaya, G., & Sogut, M. (Eds.) (2024). Proceedings Book Series-I of the ESERA 2023 Conference. Connecting Science Education with Cultural Heritage. Cappadocia, Türkiye. Nobel Bilimsel Eserler.

The appropriate APA style for referencing individual papers in the eProceedings is as follows:

[Author(s)]. (2024). [Title of article]. Sardag, M., Kaya, G., & Sogut, M. (Eds.) (2024). Proceedings Book Series-I of the ESERA 2023 Conference. Connecting Science Education with Cultural Heritage, Strand number (co-ed. [Editors of the strand chapter]), (pp. [page numbers]). Cappadocia, Türkiye. Nobel Bilimsel Eserler.

Editorial

We are pleased to present the first volume of the ESERA 2023 Conference Proceedings Series, which brings together an extensive collection of research studies under the theme of Science Teaching, Learning, and Assessment. This book features contributions from seven strands of the conference: Strand 1 *(Learning Science: Conceptual Understanding*), Strand 2 (*Learning Science: Cognitive, Affective, and Social Aspects*), Strand 4 (*Digital Resources for Science Teaching and Learning*), Strand 14 (*Evaluation and Assessment in Teaching and Learning*), Strand 17 (*Science Education in Primary and Pre-primary Learning Contexts*), Strand 18 (*Teaching and Learning Science at Middle and Secondary School (6-13 Grades)*), and Strand 19 (*Teaching and Learning Science at the University Level*).

This comprehensive volume offers a rich and diverse exploration of the research and practice shaping science education today. With a total of 56 proceedings, this book captures the latest insights from across these strands, highlighting innovative approaches, challenges, and emerging trends in the field. The studies presented delve into conceptual understanding, the cognitive and affective dimensions of learning, the integration of digital tools, effective assessment strategies, and the unique characteristics of teaching and learning across different educational levels—from early childhood to higher education.

Each section of this book is dedicated to one of the included strands, with a summarizing introduction that contextualizes the research presented in that area. This structure allows readers to navigate the diverse topics covered, while offering a cohesive view of how these strands interconnect and contribute to the broader landscape of science education.

The first two strands in this volume focus on foundational aspects of learning science: conceptual understanding and the cognitive, affective, and social dynamics that influence learning. These studies explore how students develop scientific concepts, the factors that shape their learning experiences, and the implications for teaching practices.

Strand 4 emphasizes the role of digital resources in science education, reflecting the growing importance of technology in both classroom and remote learning environments. The proceedings in this section investigate the potential of digital tools to enhance engagement, support personalized learning, and foster deeper understanding.

Evaluation and assessment are central to effective teaching and learning, as highlighted in Strand 14. This section addresses the design and implementation

of assessment strategies that support learning, promote equity, and provide meaningful feedback for students and educators alike.

Strands 17, 18, and 19 focus on the diverse learning contexts within which science education takes place, from early childhood settings to primary and secondary schools, and into the university level. The research in these sections offers valuable perspectives on the unique challenges and opportunities at each stage of the educational journey.

We are grateful to the authors, reviewers, and conference organizers whose contributions have made this volume possible. Their collective work reflects the ongoing commitment to advancing science education through rigorous research, innovative practices, and collaborative efforts. It is our hope that the studies compiled in this book will not only inform current practice but also inspire future research and development in the field.

We invite you to explore the breadth and depth of the research presented in this volume and trust that it will serve as a valuable resource for educators, researchers, and policy-makers dedicated to improving science education at all levels.

Warmest Regards,

Gokhan Kaya, Metin Sardag, Mehmet Sogut

Hacettepe STEM & Maker Lab

August 2024

Contents

Editorial **4**

Strand 1. Learning Science: Conceptual Understanding **8**
Meaningful Learning of Exoplanetology and Habitability 10
Relationship between the Different Levels of External Mediations reported by the Students and their Conceptions about the Concept of Light 18
What are High School Students' Mental Representations about Relativity Theory? 30
A Human Orrery to Help 11 Years Old Students Investigate The Shape of Planetary Orbits 42
The Literature Review on Teaching and Learning the Properties of Matter in Physics Education 52
The Effect of Simulation-Integrated Online POE Activities on Preservice Science Teachers' Conceptual Understanding of Simple Electric Circuits, Attitudes and Physics Self-Efficacy Beliefs 66
Students' Reasoning on the Dependence of Energy and Structure in Chemical Reactions 78
Investigating The International Applicability of The Maverbe Analysis Model 88
A Qualitative Model of the Evaporation Process for Science Education 100

Strand 2. Learning Science: Cognitive, Affective, and Social Aspects **108**
Research Trends in Science Education from 2017 to 2021: A Content Analysis of Publications in ESERA Conference Proceedings 110
Use of Analogies in The Teaching of Science: Post-Festum and Heuristic Analogies 118
Meeting Between Spinoza and Gilles Deleuze: From Affection to Learning in Sciences 124
The Role of Open Schooling in Community Efforts to Tackle the Silent Pandemic of Antimicrobial Resistance 132
"If you Really, Really, Want to Know it You Have to Touch it" – Researching the Role of Touch in Children's Science Education 144
Examination of Effects of Authentic Scientific Inquiry Experiences on Students' Epistemological Beliefs 158
Augmented Reality Used in Physics Experiments – A New Way to Cognitive Destress Students and Enhance Their Interest? 170
Researching the Teaching and Learning of Physics in Formal and Informal Contexts: A Study from Kindergarten to University 184

Strand 4. Digital Resources for Science Teaching and Learning **194**
Teaching Nanotechnology Concepts Through the Use of Digital Technology to Early-Primary School Children 196
Introducing the Control of Variables Strategy to Early Childhood Student Teachers With the Assistance of an Augmented Pedagogical Agent 212
Education of Elementary Teachers on The Use of Adaptive Gamification Environment in Science Education 224
Development and Implementation of A Robotic Activity Through Bbc Micro:Bit for Teaching Electricity to 5th Graders 238
Stimulating Reflection Processes with an Eye-Gaze-Augmented Retrospective in Organic Chemistry 252
The Use of Augmented Reality in Biology Subjects in the Context of Technology Integration into Education 260
Added Value of Google Lens For Identification of Marine Organisms 270
Digital Teaching Scenarios on Advanced STEM Topics for Higher Education 280
Using Augmented Reality and Simulations to Support Learning of Fundamental Electrical Concepts 290
Characterising The Educational Potentials of Citizen Science Mobile Applications Related to Biodiversity 306
The Effect of the Instructional Technologies Course Taught with The Blended Learning Model on the Pre-Service Teachers' Self-Efficacy Beliefs in Web 2.0 Practical Content Development 314
Digital Literacy in Turkish High School Science Textbooks 324
Virtual Reality: Where Art meets Science 334
(Not) Learning Alone with Augmented Reality: Continuing Results From a Comparative Study in an

Undergraduate Physics Laboratory Course 344
Digital Teaching Scenarios on Advanced STEM Topics for Higher Education 356
Using Augmented Reality and Simulations to Support Learning of Fundamental Electrical Concepts366

Strand 14. Evaluation and Assessment in Teaching and Learning 382
Development of a Scientific Practices Survey for Middle School Students 384
Assessment of Conceptual Understanding in Physics and Conceptual Tests..... 392
Formative Assessment in The Swiss Science Classroom: What Support Do Textbooks Offer?..... 400
Assessment of Science Competencies in Danish Primary and Lower Secondary Education 406
Investigating Gender Differences in Confidence and Accuracy in the Force Concept Inventory (FCI)416
Characteristics of Written Examination Questions That Differentiate Between Students Who Have Completed Practical Work in Hands-On and Other Ways in Science Lessons 422
Prior Knowledge in Physics: The Distribution Within Different Subgroups of Students Before Attending Physics Minors 430

Strand 17. Science Education in Primary and Pre-primary Learning Contexts440
Are Games Effective to Learn and Think about the Nature of Science in Primary Education? 442
Consideration of the Impact of a Professional Learning Program to Support the Teaching and Assessment of Practical Primary Science in England 456
Plant and Animal Photographs in the Science and Environment Textbooks of the Greek Primary School 464
“It’s Only a Drawing!” - Learning to Model and About Models in Primary School 472
A Teaching Innovation Proposal to Reduce Plant Blindness in Early Childhood Education..... 486
Early STEM Education: A Catalyst for Science Literacy in Preschool Children 492

Strand 18. Teaching and Learning Science at Middle and Secondary School (6-13 Grades).....504
Teacher Questions in Secondary Biology Classrooms: Making Scientific Practices Meaningful for Students 506
A Conceptual Framework for Digital Research Skills in Secondary Science Education..... 518
Secondary Students Drawing Comics for Physics Learning..... 524
Semantic Waves for Revealing Different Levels of Knowledge in Middle-School Science Classes ... 530
Digital Research Skills: Application in Secondary Science Education..... 540
Evaluating A Teaching-Learning Sequence (TLS) About Acid-Base Reactions In Upper Secondary School 546
Language And Middle-School Science: What Do Children And Teachers Gain By Implementing Semantic Waves In Teaching? 560
Writing Fiction to Expose Learning. An Adventure with Genetics and Evolution 576
Designing an Educational Escape Room: Evidence-Based Guidelines and Practical Advice 586

Strand 19. Teaching and Learning Science at the University Level.....596
Culturally Responsive Methods in STEM Education 598
Understanding Professional Learning in Higher Engineering Education through Signature Pedagogies 612
Navigating The STEM University Transition: Effects of a Combined Mindset and Learning Strategy Online Intervention..... 620
Enhancing STEM Teaching in Higher Education: A Study of Project-Based Pedagogical Development for Novice Teaching Assistants 634
STEM Based Astronomy Activities with Ancient Minoan Analogue Computer 647



Strand 1

Learning Science: Conceptual Understanding

Ana Sofia Afonso & Massimiliano Malgieri

Strand Chairs/Co-editors

Foreword

This chapter, titled Learning Science: Conceptual Understanding, is part of the ESERA 2023 Conference Proceedings and comprises seven insightful studies that delve into one of the core challenges of science education—how learners develop and refine their understanding of scientific concepts. Conceptual understanding is at the heart of meaningful science learning, as it enables students not only to grasp fundamental principles but also to apply them in diverse contexts and connect them with real-world phenomena.

The proceedings included in this chapter explore a range of topics central to conceptual understanding, including cognitive processes, instructional strategies, and the role of prior knowledge in learning. The research examines both the challenges and opportunities in fostering deep understanding of key scientific ideas, addressing questions such as: How do learners build robust mental models of scientific concepts? What are the most effective ways to address misconceptions? How can educators design learning experiences that promote long-term retention and transfer of scientific knowledge?

These studies are diverse in both their methodologies and their educational contexts, offering valuable insights for educators, researchers, and practitioners. From exploring the impact of specific teaching interventions to investigating the conceptual pathways learners take, this chapter presents findings that can inform better practices in science education and help bridge the gap between what students are taught and what they truly understand.

The contributions in this chapter highlight the importance of integrating cognitive science principles with pedagogical practices to improve conceptual understanding. By exploring both the successes and the ongoing challenges in this area, these proceedings underscore the need for continued research and innovation in teaching strategies that support conceptual learning.

We are grateful to the authors and reviewers whose work has made this chapter possible. Their contributions advance our understanding of how learners make sense of scientific concepts and provide actionable insights that can improve teaching and learning outcomes.

We invite you to engage with the research presented in this chapter and hope that the findings will inspire further exploration and application in your own educational practice and research endeavors.

Meaningful Learning of Exoplanetology and Habitability

Gian Michaelsen[1], Agostinho Serrano[2]

[1] PhD Student Universidade Luterana do Brasil, Brazil,

[2]PhD Teacher and Advisor Universidade Luterana do Brasil, Brazil

This research aimed to effectively incorporate Exoplanetology and Habitability concepts into high school education through a process of didactic transposition. The primary objective was to demonstrate a deep transformation in students' knowledge and understanding. To achieve this, a comprehensive didactic guide was developed, enabling students to engage in the analysis of real exoplanet data and make informed inferences about their habitability. The study employed a multifaceted assessment approach, including pretests, posttests, and specially designed interviews to address novel problem situations. The data obtained from these student activities were meticulously analyzed using Textual Analysis as outlined by Laurance Bardin. This analysis provided clear evidence of the maximum transformation of knowledge among students in various problem-solving scenarios. Such a transformation is a key indicator of meaningful learning, as posited by established educational theories. This research not only highlights the successful integration of advanced scientific concepts into high school curricula but also underscores the effectiveness of meaningful learning in fostering deep understanding and application of knowledge in different contexts.

Keywords: Physics Teaching, Meaningful Learning, Didactic Transposition

Introduction

The impetus for developing the instruction of Exoplanetology and Habitability arose from observing the frequent coverage of these topics in the media. However, a review of literature revealed a scarcity of articles addressing their pedagogy. This gap is likely attributable to the relative novelty of the research field. As Michaelsen (2022) notes, since the discovery of the first exoplanet in 1992, over 5,000 exoplanets have been identified. The Planetary Habitability Laboratory (LHP) at the University of Puerto Rico classifies 55 of these as potentially life-sustaining. These findings suggest that incorporating this subject into educational curricula could significantly engage and stimulate student curiosity, an essential precursor to meaningful learning.

Awakening student curiosity is vital for fostering a readiness to learn, which is crucial for the occurrence of meaningful learning. As defined by Ausubel (1968), meaningful learning is enduring and is less likely to be forgotten compared to rote learning, which is transient and prone to being completely forgotten. Ausubel posits that the epitome of meaningful learning is evidenced by the profound transformation of knowledge, wherein a student effectively synthesizes their understanding and applies it adeptly in novel contexts, distinct from any previously encountered examples.

Given this backdrop, the central inquiry of our study is: How can the didactic transposition of Exoplanetology and Habitability concepts to high school students efficaciously promote meaningful learning, as demonstrated by the profound transformation of knowledge through student engagement with authentic exoplanet data and complex problem-solving scenarios? This investigation aims to assess the potential of integrating these emerging scientific topics into the curriculum. By evaluating their effectiveness in sparking student curiosity and learning readiness, the study endeavors to elucidate how these innovative and compelling subjects can culminate in deep, meaningful learning experiences. Therefore, the research examines the process by which students assimilate these advanced concepts and apply them inventively in challenging and unfamiliar situations. In light of the burgeoning interest in exoplanetary sciences, it is imperative to acknowledge the historical journey of exoplanet discovery and its profound impact on our understanding of the cosmos. This enhanced perspective not only enriches the field of astronomy but also offers valuable insights into the interdisciplinary nature of scientific inquiry, laying a foundation for future explorations and educational endeavors.

Meaningful Learning

Ausubel, Novak, and Hanesian's (1968) theory of teaching and learning proposes that concepts can be assimilated in two primary ways: mechanically and meaningfully. Mechanical learning involves the student absorbing information without fully comprehending it, which restricts their ability to apply concepts in new situations and engage in critical reflection. This approach results in a passive learning style, where the student is merely a recipient of information rather than an active participant in their knowledge construction.

Conversely, in Meaningful Learning Theory (MLT), as Ausubel describes, concepts are retained more durably in the student's memory. Here, the student is an active agent in constructing knowledge. Meaningful learning hinges on opportunities for students to reflect and ascribe meaning to knowledge in a purposeful manner.

Meaningful Learning Theory (MLT) underscores the significance of prior knowledge in the learning process. Valadares (2011) asserts that meaningful

learning necessitates a deliberate and relevant connection between new information and pre-existing knowledge within the student's cognitive structure. Ausubel originally introduced the concept of "subsumers", which are pivotal concepts in a student's cognitive structure, essential for meaningful learning. The assimilation of new concepts occurs through these subsumers.

Additionally, it is crucial for didactic materials to logically present concepts, enabling students to connect them with their subsumers. The student's predisposition to learn is another critical element for meaningful learning. Santos (2007) notes that this predisposition can be enhanced through positive learning experiences, contributing to the development of MLT.

We deduce that curiosity serves as a key factor in fostering a student's readiness to learn, crucial for the effective assimilation and interpretation of new concepts. Specifically, by stimulating curiosity about exoplanets and their habitability, we encourage meaningful learning. This approach not only prepares the student for this type of learning but also enriches and transforms their cognitive structure. Consequently, the complexity of this structure is augmented, facilitating the integration of new concepts in a more profound and significant manner. The following session will detail how this was accomplished.

Methodology

This research was conducted during the COVID-19 pandemic; therefore, our initial idea was modified to align with the constraints of that period. Given the impracticality of face-to-face classes, we developed a didactic guide – a document providing the necessary theoretical foundation for contextualizing the activity, alongside qualitative and quantitative analyses of real exoplanet data.

Our literature review identified two pivotal works for the research: LoPresto and Ochoa (2017) and Della-Rose et al. (2018). These studies inspired our didactic development. LoPresto and Ochoa's work detailed a process for identifying potentially habitable exoplanets using real data, while Della-Rose and colleagues developed an app to help students analyze exoplanet data.

Inspired by these studies, our didactic guide offered both theoretical and practical insights. Initially, the guide introduced concepts in a style and language suitable for high school students, setting the stage for understanding exoplanet habitability. Following this, students engaged in qualitative data analysis, applying methods adapted from Della-Rose et al. (2018) to determine specific details of the studied exoplanets. In the classification phase, students used the calculated real data to categorize the exoplanet, comparing it with planets in the solar system. During the inference stage, students utilized parameters from LoPresto and Ochoa (2017) to deduce the habitability of their assigned exoplanet. This stage marked the first

evidence of significant knowledge transformation, with students articulating and reasoning their inferences.

In addition to the didactic guide, we employed a pretest, posttest, and individual interviews to assess student comprehension and development. The pre and posttests comprised three open-ended questions, allowing students to freely express their knowledge and beliefs. The interviews, structured following Gil's methodology (2006), included 11 questions focused on comparing pre and posttest results and identifying the extent of knowledge transformation.

The study involved 19 high school seniors from a public school in southern Brazil. Using Google Meet, we conducted the activities in group sessions, while interviews were conducted individually. The qualitative data from these activities and interviews were analyzed using Laurence Bardin's Content Analysis method (2016). Bardin (2016) describes content analysis as a versatile set of techniques for identifying commonalities in responses and categorizing them based on frequency.

This article presents results concerning the profound transformation of knowledge in two problem scenarios posed during the interviews:

1. "What impact would a halving of the orbiting star's temperature have on your studied exoplanet, and how might this change affect its habitability?"
2. "Could life potentially develop on the exoplanet under these altered conditions?"

These questions required students to integrate the content learned through the didactic guide, demonstrating their ability to apply theoretical concepts to practical scenarios.

The methodological approach of this research is rooted in the principles of didactic transposition, carefully chosen to bridge the gap between complex scientific concepts and high school comprehension. This section delves into the rationale behind this pedagogical choice, offering insights into its effectiveness in demystifying advanced scientific ideas for young learners.

Results

Through the content analysis performed (Bardin, 2016) we classified the responses of the two problem situations into 5 different categories. The first being "Corrects", when the student's response was adequate, clear and precise; "Correct (insecure)" when the answer was adequate but the student was not sure; "Insecure" when the student's position is not clear; "Wrong (insecure)" when the student was not sure about an inadequate answer and "Wrong" for when the student was sure about a wrong answer.

The solutions for the first problem situation had a “Correct “ index of 52.6%, “ Correct (insecure)” of 42.1%, “ Insecure “ and “ Wrong (insecure)” with 0% and “ Wrong “ with 5,2%. In the second problem situation, the rate of “Correct” solutions was 10.5%; “Correct (insecure)” of 52.6%; “Insecure” of 26.3%; “Wrong (insecure)” with 10.5% and “Wrong” with 0%.

We also identified the conceptual relationships built by the students and classified them into the primary categories “Concepts related correctly” and “concepts related wrongly and/or without foundation.”

Table 1. Concept relations made by the students.

	Conceptual relationship	Frequency	Percentage
Correctly related concepts	Star temperature and planet temperature	18	50
	Star temperature and Habitable Zone	7	19.4
	Star temperature and star type	2	5.5
	Planet temperature and possibility	2	5.5
	Planet temperature and physical state of water	5	13.8
	Planet temperature and planet atmosphere	2	5.5
Incorrectly Related and/ or Unjustified Concepts	Star temperature and luminosity	1	12.5
	Star temperature distance from planet	1	12.5
	Star temperature and planet mass	2	25
	Star temperature and planet type	1	12.5
	Star temperature and planet temperature	1	12.5
	Star temperature and planet density	1	12.5
	Star temperature and planet atmosphere	1	12.5

An in-depth examination of student responses reveals significant insights into their comprehension and application of the learned concepts. This analysis not only sheds light on the areas where students excelled but also identifies common misconceptions and challenges, offering a comprehensive overview of the learning outcomes.

Discussion and Conclusion

The analysis of the first problem situation revealed that the majority of students, over 90%, provided responses close to being correct, albeit sometimes uncertain. This high percentage indicates that nearly all students successfully applied newly acquired knowledge to the given problem. In contrast, for the second problem situation, the accuracy rate was lower. However, most responses were classified as ‘correct, albeit uncertain’.

Comparing these two scenarios, it becomes evident that the second required more critical and analytical thinking from the students. While the first question might have seemed “easier” in relating the star’s temperature to other exoplanet aspects,

considering the possibility of life necessitated a more complex understanding. Despite uncertainties, students demonstrated the ability to develop coherent reasoning for the scenario. This suggests that the interview setting might have induced nervousness, affecting their confidence in responses.

The success in achieving the highest level of knowledge transformation is also evident in the analysis of the conceptual relationships established by students. For instance, student B02 demonstrated a thoughtful process of relating concepts: "So let me think a little. Half. I think it would affect a lot. He for reasons of. It is to exchange a lot because the temperature is very high. ... The habitable zone would change for sure because his temperature changed a lot. The density of it wasn't going to change...You're going to see the life in it because I still don't believe so. So that if it would be more likely he would be less far from being able to have life. Because, as I said, one of the main factors of his not being able to have life was his extremely high temperature. So, there would be more possibility of him having. Life even he has water itself in liquid state. I don't know if it would be possible right now upside down like that, but.... I think it was pretty hot, I think maybe even half. It was pretty hot. But I think that would be just that....". In B02's response, the conceptual relationships are appropriately and coherently connected.

Upon analyzing all students' conceptual relationships, 36 were found to be correctly established, while only 8 were incorrect or unfounded. In some instances, these relationships were formulated during the inference stage of the exoplanet's habitability in the creation of the didactic guide. The adept application of these concepts in new and challenging scenarios underscores the students' conceptual mastery.

The findings of this study indicate the successful integration of Exoplanetology and Habitability concepts in high school education via the didactic guide. This integration was not only successful but also demonstrated evidence of meaningful learning, as shown by the significant knowledge transformation observed in many students. This research highlights that teaching emerging sciences can greatly appeal to students, a vital aspect for meaningful learning across various subjects and knowledge areas. As we conclude, this research highlights the long-term impact of integrating complex scientific concepts into high school education. It underscores the role of meaningful learning in preparing students for future academic and professional challenges, emphasizing the need for continual exploration and integration of emerging scientific fields into educational curricula. Future research should focus on expanding the application of the didactic guide across diverse educational settings and cultural contexts. Such studies could provide invaluable insights into the universal applicability and adaptability of advanced scientific concepts in secondary education. Additionally, exploring alternative and innovative approaches to

science education at the high school level, particularly those that emphasize meaningful learning and practical application, could further enrich the field and offer new perspectives on educational methodologies.

References

Ausubel, D. P., Novak, J. D., & Hanesian, H. (1968). Educational psychology: A cognitive view. Holt, Rinehart & Winston.

Bardin, L. (2016). Análise de conteúdo. Edições 70.

Della-Rose, D., Carlson, R., Harpe, K., Novotny, S., & Polsgrove, D. (2018). Exoplanet science in the classroom: Learning activities for an introductory physics course. *The Physics Teacher*, 56, 170. https://doi.org/10.1119/1.5025299

Gil, A. C. (2006). Métodos e técnicas de pesquisa social (5th ed.). *Atlas*.

Michaelsen, G. A., & de Andrade Neto, A. S. (2021). Teaching of physics in the context of planetology and habitability in high school: A case of successful online meaningful learning. *RENOTE*, 19(2), 437-446.

Moreira, M. A. (2012). O que é afinal aprendizagem significativa? Instituto de Física da UFRGS. http://moreira.if.ufrgs.br/oqueeafinal.pdf

Lopresto, M., & Ochoa, H. (2017). Searching for potentially habitable extra solar planets: A directed-study using real data from the NASA Kepler-Mission. *Physics Education*, 52, 065016. https://doi.org/10.1088/1361-6552/aa8740

Rocha, A. M. (2019). Astronomy teaching in the context of exoplanet discoveries. Graduate Program in Teaching Natural Sciences and Mathematics: UFRN.

Santos, F. M. T. dos. (2007). As emoções nas interações e a aprendizagem significativa. *Ensaio Pesquisa em Educação em Ciências* (Belo Horizonte), 9(2), 173-187. Retrieved May 1, 2022, from https://www.scielo.br/j/epec/a/jVJt79Q5yXpjfyWGD3BrJKs/?format=pdf&lang=pt

Valadares, J. (2011). A teoria da aprendizagem significativa como teoria construtivista. Aprendizagem Significativa em Revista, *1*(1), 36-57.

Relationship between the Different Levels of External Mediations reported by the Students and their Conceptions about the Concept of Light

dos Anjos, Juliana Rodrigues, Serrano, Agostinho, de Souza, Maira Giovana

Post-Graduate Program in Science and Mathematics Teaching – Lutheran University of Brazil, Brazil

This study aimed to investigate the relationship between the level of mediation used by high school students and their conceptions about the concept of light. This link is analyzed through the "External Mediation Level Profile", a tool that analyzes the representations used by the students' verbal and gestural languages, decoded after individual interviews, under the perspective of Cognitive Networks Mediation Theory (CNMT). For the CNMT, there are four different levels of external mediation - Psychophysical, Social, Cultural and Hypercultural - that facilitate the development of mental images by students when answering questions about concepts relevant to light. By raising the profile of each student and confronting him with their conceptions of the nature of light, it is possible to understand their relationships. Results indicated that the level of mediation used by students was closely related to their understanding of the concept of light. The study highlights the importance of utilizing various forms of external mediation in the classroom to aid in students' understanding and retention of scientific concepts.

Keywords: Students' conception; behavior of light; external mediation level profile.

Introduction

The study of spectroscopy has played a crucial role in advancing scientific knowledge (Filgueiras, 1996) and its themes are highly relevant in the context of basic education. The subjects, particularly those related to the concept of light, are essential for a deeper understanding of the world today. Studies have shown that students lack an adequate understanding of the concept of light (Schützenhöfer, 2017) even after taking an optics class. The aim of this research was to develop activities related to spectroscopy for high school students using various resources such as simulations, videos, texts, mobile applications, and presentations.

In this work, we present results from the interaction of students with different extracerebral tools for the study of concepts related to the behavior of light and

matter. The research used the Cognitive Networks Mediation Theory (CNMT) as its theoretical framework (Souza et al, 2012). According to this theory, the human brain uses external resources (extracerebral tools) to expand its information processing capacity. Interaction with external resources essentially occurs through four different levels of mediation: psychophysical, with the use of physical objects, stimulating sensory-motor schemes; social, through interaction between individuals in a group; cultural, using symbolic schemes such as language; and hypercultural, which occurs through technological tools.

It is natural that the student, when solving any problem, uses not just one, but several different external mediations that intertwine and make up the repertoire of mental images that the individual uses to solve the problem. For this purpose, we built the External Mediation Level Profile (EMLP) for each participant on a given concept, which indicates which mediation levels are preferably used by the student in solving a specific problem. The idea of creating an EMLP was originally inspired by the constructs of Bachelard (1985) and Mortimer (1995), who studied the problem of solving scientific problems with the Epistemological and Conceptual Profiles, respectively. The EMLP was developed based on the four different levels of CNMT mediation described above, for each student's concept of Light and, in this work, it is represented by group of students. This profile was useful to visualize the most relevant mediations, as well as the transitions between external mechanisms.

The different objects or images used in the classroom (Kesonen, Asikainen, & Hirvonen, 2017), as well as the different experiences of each studehipent (Uzun, Alev, & Karal, 2013) influence the understanding of the concept of light. Thus, it is possible to understand that different mechanisms of external processing can help in the cognition and production of mental images in the students, in different levels of mediation, influencing the students' conceptions. The activities were developed with high school students, with their informed consent, and the guiding problem of the research revolves around the following question: *What is the relationship between the level of mediation preferably used by students and their conceptions about the concept of light?*

Methodology

The research was conducted with 35 third-year high school students from various public schools in Brazil in a remote environment. The students were given tasks such as presentations, videos, simulations, and tests, including the challenge of creating a virtual experimental activity - building a spectrophotometer using a mobile application, SpectraUPB[1], and low-cost materials. The students completed pre-test and post-test questionnaires immediately before and after all activities. Based on the results of the questionnaires and the progress of the

1 Available in: https://www.upb.edu/en/contenido/spectrometry-software-for-android

research, 12 students were selected for an interview.

The interviews were conducted using the Report Aloud protocol (Trevisan et al., 2019), with questionnaires guided by the teacher-researcher individually. All interviews were recorded on video and transcribed for analysis using Descriptive Gesture Analysis (Monaghan & Clement, 1999), aiming to establish a connection between the descriptive gestures performed by the students and their mental images, as well as investigating the origin of these images.

In the transcriptions, the students' names are not used; they will be referred to by the letter "S" followed by a number, such as S1 (Student 1), and so on for the other students. The gestures produced by the students are identified within square brackets "[hashtag]" (#) and letters corresponding to the mental image identified through the movement or position of each participant's hands.

Results and Discussions

With the data obtained, we identified the mediations reported by the students and their conceptions about the concept of light, categorizing the students based on models. Three categories were created from the explanations of each student throughout the interview: (I) Wave model; (II) Ray or beam model, and (III) Hybrid model. A purely of particle model was not evident in the students' speech, so it does not appear in this analysis. The categorical system presented here does not take into account the historical evolution of models regarding the nature of light or the inconsistencies present in students' conceptions regarding scientific models of light. The explanation for each identified model is provided in the table below.

The construct that allows identifying which external processing mechanisms are most used to create mental images is the External Mediation Level Profile. From the analysis of the mediations, three groups emerged naturally: the group of students in which the mental images are, predominantly, originating from Psychophysical Mediation (P), the one from Hypercultural Mediation (H), and the group with preference for Cultural Mediation (C). It is not within the scope of this article to analyze the origin of mediations and the construction of profiles, but rather to understand that, within the group of students analyzed, there is a relationship between the most frequently used mechanisms and their images regarding the concept of light.

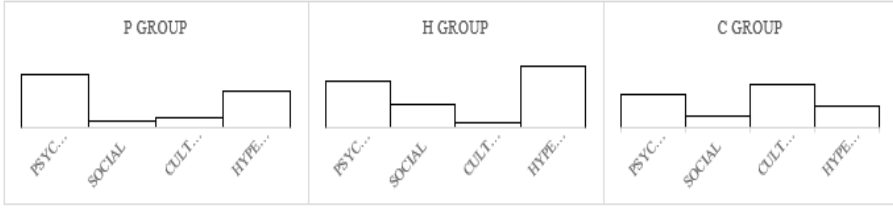
Table 1. Categorical System of Light Models.

Models	Explanation
(I) Wave model	Light as a wave. In this category, students who identify and explain certain phenomena related to light while conceptualizing it as a wave are included. Some students, when referring to the wave, associate its mode of propagation with wave disturbance. This category may encompass students who discuss and explain characteristics of the wave-like behavior of light or those who do not.
(II) Ray or beam model	Light as "something" emitted by a luminous source, a copy of reality. In this category, students may be included who discuss certain characteristics of light, such as its speed, dispersion of white light, reflection, or refraction, without providing elaborate explanations and visualized in the form of rays, beams, or lines.
(III) Hybrid model	Students, when sharing their understanding of the behavior of light, mention a duality without recognizing any distinction between the wave and corpuscular models, simultaneously associating elements of both wave and particle. Alternatively, at one moment, the student explains the phenomenon by envisioning particles, and at another moment, imagines waves, indicating a connection between the models, whether articulated or not.

Source: The survey (2023).

The organization of the analyses is found below and is separated by group, based on the three main questions of the pre-test and post-test, carried out during the interview. The first being related to the image of "what is light" (1), the second with the idea of "luminous emission" (2), and the third about what would be "the absorption of light" (3). For each question, we selected only one student representing one of the three emerging groups (Groups P, H, and C) and employing one of the models in their explanations. For instance, Group P is represented by student S1. In the analysis presentation, responses surrounding one of the main questions (in this case, question 1) will be reported, highlighting the externally mediated processing mechanisms identified through the External Mediation Level Profile (figure 1) and their conception of the concept of light.

Figure 1. External Mediation Level Profile of Groups P, H and C, respectively.



Source: The survey (2023).

Table 2. Psychophysical Group. Each "S" represents a "student" and for each question, a predominant Model was identified.

	S1	S2	S6	S13
Question 1	(II) Ray or beam model	(II) Ray or beam model	(II) Ray or beam model	(I) Wave model
Question 2	(II) Ray or beam model	(II) Ray or beam model	(II) Ray or beam model	(II) Ray or beam model
Question 3	(II) Ray or beam model	(II) Ray or beam model	(II) Ray or beam model	(II) Ray or beam model

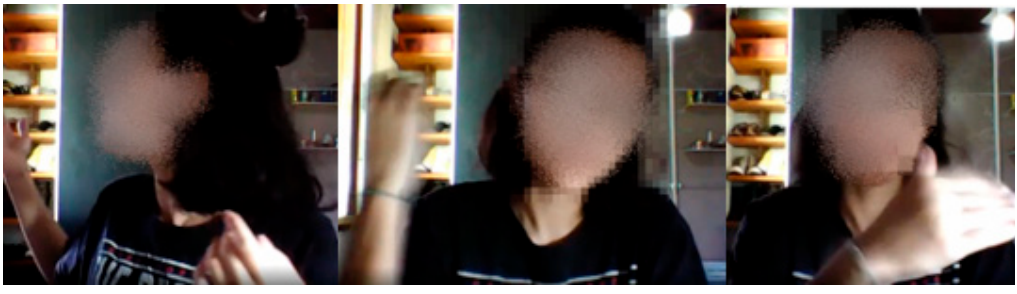
Source: The survey (2023).

For the group in which psychophysical mediation was most used, the Ray or Beam Model stood out. In this case, light is seen as "something" that is emitted by a source, a copy of reality. In this category were included students who talk about some characteristics of light, without much explanation, and with visualization in the form of a ray, beam, or lines.

We analyzed the student's conception of light behavior to explain certain phenomena, following the interview protocol, identifying what the student imagines. Student S1, belonging to Group P, when talking about light, states:

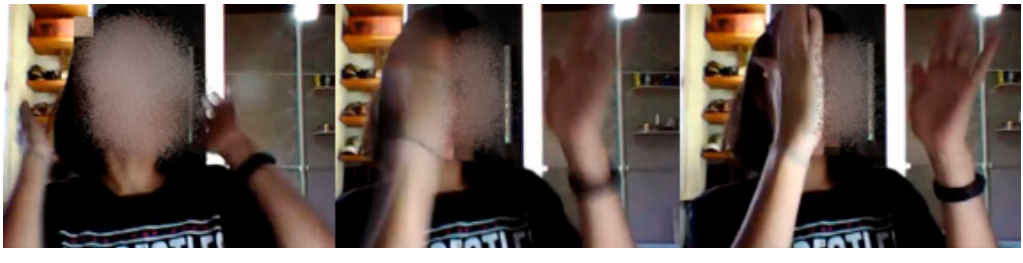
> **S1:** I end up imagining it as something straight [#RA], like when I imagined my window and the ray coming in directly to here. But of course, they are waves, I understand that. I just end up imagining it as if it were a spotlight [#HO].

Figure 2. The sequence of images indicates a forward movement with one hand, showing how a luminous ray would be, in a straight line. Dynamic image. #RA



Source: The survey (2023).

Figure 3. The sequence of images indicates a forward movement with open hands, as if delimiting the size of the light beam from a spotlight. Dynamic image. #HO



Source: The survey (2023).

It can be observed that she uses elements of light interaction with her environment to formulate her response, characterizing a psychophysical mediation. Even though she understands that they are waves (according to her), S1 ends up imagining it as "something straight." In this case, student S1 fits into Model II, of rays, beams, or lines, as she uses (imagines) this idea to explain phenomena throughout the interview.

Table 3. Hypercultural Group. Each "S" represents a "student" and for each question, a predominant Model was identified.

	S3	S4	S5	S7	S10	**S11**	S12
Question 1	(III) Hybrid model	(I) Wave model	(III) Hybrid model	(I) Wave model	(II) Ray or beam model	(I) Wave model	(I) Wave model
Question 2	(I) Wave model	(I) Wave model	(II) Ray or beam model	(III) Hybrid model	(II) Ray or beam model	(I) Wave model	(III) Hybrid model
Question 3	(I) Wave model	(III) Hybrid model	(II) Ray or beam model	(I) Wave model	(II) Ray or beam model	(I) Wave model	(I) Wave model

Source: The survey (2023).

It is noticed that, for the Hypercultural Group, the Model that stands out among the questions is the (I) Model Wave. This category includes students who point out and explain some phenomena related to light, imagining it as a wave. Some students, when referring to the wave, associate its form of propagation to a wave disturbance. Students who speak and explain or not the characteristics of the wave behavior of light are included.

Observing the responses of student S11 to the third question involving the concept of light absorption, it is possible to note the strong presence of the "wave" concept. Even though the student uses the idea of light interacting with objects in the surrounding environment, their explanations are related to imperceptible physical phenomena, as stated:

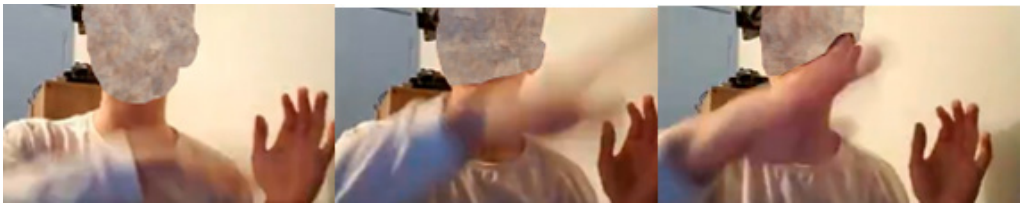
> **S11:** Now I'm wearing a white shirt, so when the light comes to it, it absorbs and emits all colors, doesn't retain the wavelength, tends to return the colors that, when combined in your eye, will appear white. And if it were a red shirt, the white wave would come, with the sum of all the colors of the visible spectrum, and the atoms of the shirt would energize, and when they emit a wave, they would still contain the energy of the other colors, wavelengths. Like in simulations.

The student is then asked to explain what they imagine about light absorption, describing the experiment they conducted at home:

> **S11:** Here is the tube [#TU], here is the glass [#C], and here is the light emission [#EL]. The white wave, right, has all the colors, as we've already mentioned, from the visible spectrum, and in the glass, because it looks blue to us, it's because that white light, when it hits there, that

liquid absorbs all the colors, keeps them, except for the blue that it will emit along with other nearby wavelengths. That's why when we visualize it, we only see blue. It was initially white, it will pass through the blue, pass through the CD, and reach the cellphone with only blue, which was expected due to the filter that the glass becomes.

Figure 4. The sequence of images illustrates a movement with one of the arms moving up and down, with the palm of the hand facing downward, indicating a wave. Dynamic image. This sequence of images is labeled with the hashtag #ON.



Source: The survey (2023).

Student S11 constructs their explanations based on a wave, and despite using the atom to explain the occurrence of phenomena, which could indicate a Hybrid Model, when mentioning light or colors, they reference the Wave Model (I). For S11, the interaction of a wave with the atom (the electron of an atom) leads to electronic transitions, and the light (or color) resulting from this relationship is observed as a wave. In the end, he reinforces this thought by saying:

> **S11:** I still have a greater tendency to imagine it as a wave; maybe by studying more modern physics, perhaps quantum physics, it will become simpler to think about the photon. For now, I still see it as an electromagnetic wave. Like X-rays, infrared, ultraviolet that scientists use to see footprints and everything else, it's still something that seems more present, at least in my mind that leans more towards this interpretation.

Finally, there is the Cultural Group, represented by only one student (S9). Learners allude to a duality without acknowledging a clear distinction between the wave model and the corpuscular model, concurrently linking elements of waves and particles. Alternatively, a student elucidates the phenomenon by envisioning particles, while at another instance, they contemplate waves, demonstrating an interrelation between the two models.

Table 4. Cultural Group. For each question, a predominant Model was identified.

	S9
Question 1	(III) Hybrid model
Question 2	(III) Hybrid model
Question 3	(II) Ray or beam model

Source: The survey (2023).

During the interview, the student S9 explains that they already had knowledge about light sources of different colors and, in recounting this, mentions characteristics such as the wavelength of light. Continuing in the same conversation, S9 explains how they imagine this wavelength, making some gestures. Furthermore, they relate the wavelength to frequency and, consequently, to the energy of a photon.

> **S9:** I already knew, from what was seen, that a light source, a red lamp or LED, would emit only a beam of light with a characteristic wavelength of red, while a blue LED would emit a characteristic wavelength of blue, as observed in the experiment.

> **S9:** I think we have a wave [#ON] and that the wavelength is the distance between two consecutive points [#CO], which will repeat in the oscillatory motion [#ON]. Moreover, the longer the wavelength, the lower the frequency, which, in turn, is proportional to the energy of the photon [#FT] that describes the motion, as shown in the equation: E = hv.

Figure 5. The sequence of images indicates a movement with the index finger, moving it up and down, demonstrating oscillation. Dynamic image. #ON



Source: The survey (2023).

Figure 6. The image indicates a position between the thumb and forefinger, exemplifying a sphere. Static image. #FT



Source: The survey (2023).

Regarding the understanding presented by S9 in the face of the observed phenomena, it is evident that the student blends different models of light when illustrating what was being imagined. He complements his reasoning with the image of light that, for him, would be the most acceptable.

> **S9:** I imagine one and the other at the same time. I think the most acceptable is a particle with a very negligible mass that behaves like a wave, performing an undulatory motion [#ON]. And because of that, it will present characteristics of both.

It is important to note that most of the images produced by S9 originated from a book, evidence of cultural mediation. When talking about light emission, he envisions the idea of wavelength, gesturing an undulatory movement, and despite mentioning the light beams that come from different LEDs, he mentions the equation of the energy of a photon. In other words, he uses a beam model to demonstrate what he was imagining, without abandoning the ideas of particles and waves, as seen in different moments. Therefore, the student presents the conception of a Hybrid Model (III) for light.

For the results section, we selected three students, S1, S9, and S11, aiming to represent each Group (P, C, and H) and each Model (II, III, and I). In the analysis of the interviews of the three students, we obtained a total of 64 descriptive gestures.

Conclusion

This contribution investigated the production of mental images by participants throughout the interviews and their relationships with models regarding the behavior of light. Various mediations and their transitions were displayed based on the External Mediation Level Profile. It was possible to observe each student's conceptions of the concept of light when responding to the main investigation questions.

We identified that students in Groups H or C more frequently mention terms such as "photons," "atom," "waves," "particles," or "waves with particles," and also "particle with undulatory movement," along with more detailed descriptions of the wave characteristics of light, something not identified in students from Group P. The latter base their ideas on images of "beam/ray" (alluding to geometric optics and the visual interaction of light with objects) for light. In other words, depending on the level of mediation used, students were able to explain their thoughts using a broader repertoire of concepts/representations. The production of mental images in a student's cognitive structure indicates that there is a process being imagined when explaining a phenomenon.

The origin of these images depends on external mediation, which, in turn, is connected to each individual's experiences, also highlighted in the literature review. Finally, the study will be used as a reference in future research since the results highlighted that the predominant type of mediation used in creating students' images can determine their conceptions of the behavior of light.

Acknowledgement

This work was supported by the Coordination for the Improvement of Higher Education Personnel (CAPES), under grant 88887.661734/2022-00.

References

Bachelard, G. (1985). The New Scientific Spirit. Rio de Janeiro: Tempo Brasileiro.

Filgueiras, C. A. (1996). Spectroscopy and Chemistry – From the Discovery of New Elements to the Threshold of Quantum Theory. Química Nova na Escola, 3, 22-25. Retrieved from http://qnesc.sbq.org.br/online/qnesc03/historia.pdf

Kesonen, M. H. P., Asikainen, M. A., & Hirvonen, P. E. (2017). Light Source Matters–Students' Explanations about the Behavior of Light When Different Light Sources are used in Task Assignments of Optics. Eurasia Journal of Mathematics, Science and Technology Education, 13(6), 2777-2803. https://doi.org/10.12973/eurasia.2017.01253a

Monaghan, J. M., & Clement, J. (1999). Use of a computer simulation to develop mental simulations for understanding relative motion concepts. International Journal of Science Education, 21(9), 921-944. https://doi.org/10.1080/095006999290237

Mortimer, E. F. (1995). Conceptual Change or Conceptual Profile Change? Science & Education, 4 (1), 267-285. https://doi.org/10.1007/BF00486624

Schützenhöfer, C. (2017). Students' conceptions on white light and implications for teaching and learning about colour. Physics Education, 52(4), 044003. https://doi.org/10.1088/1361-6552/aa6d9c

Souza, B. C.; Silva, A. S.; Silva, A. M.; Roazzi, A.; Carrilho, S. L. S. (2012). Putting the Cognitive Mediation Networks Theory to the test: Evaluation of a framework for understanding the digital age. Computers in Human Behavior, 28 (1), 2320-2330.

https://doi.org/10.1016/j.chb.2012.07.002

Trevisan, R., Serrano, A., Wolff, J., & Ramos, A. (2019). Peeking into students' mental imagery: the Report Aloud technique in science education research. Ciência & Educação, 25(3), 647-664. https://doi.org/10.1590/1516-731320190030004.

Uzun, S., Alev, N., & Karal, I. S. (2013). A cross-age study of an understanding of light and sight concepts in physics. Science Education International, 24(2), 129-149. Retrieved from https://files.eric.ed.gov/fulltext/EJ1015829.pdf

What are High School Students' Mental Representations about Relativity Theory?

de Souza, Maira Giovana, Serrano, Agostinho, dos Anjos, Juliana Rodrigues

Post-Graduate Program in Science and Mathematics Teaching – Lutheran University of Brazil, Brazil

Considering the significance of approaches related to Modern Physics in basic education, this research explores the outcomes following the implementation of activities with high school students that incorporate both Special and General Relativity Theory. Given the inherently abstract nature of this theory, fostering the development of mental representations among students becomes a valuable tool for enhancing comprehension. Consequently, an analysis was conducted through video recordings of students with the goal of identifying the mental representations employed in the comprehension of Relativity. The findings revealed evidence of the use of both verbal (propositional) and nonverbal (imagistic) mental representations in various situations. These results align with the Dual Coding Theory, as students employed both coding systems in their learning process.

Keywords: Mental representations, relativity, dual coding.

Introduction

The fundamental goal of basic education is to prepare students with the ability to comprehend the contemporary world and foster civic awareness. In this context, Modern and Contemporary Physics (MCP) plays an increasingly pivotal role in elucidating the developments of the 20th and 21st centuries, explaining emerging technologies (Boublil et al., 2023).

Einstein's Relativity Theory (RT) holds particular appeal for students in this domain (Hartle, 2006; Henriksen et al., 2014; Kersting, 2019). The advancement of RT has not only contributed to the development of widely utilized technologies, such as the Global Positioning System (GPS), but has also gained visibility through recent detections of gravitational waves and images of black holes, drawing attention on social media and bringing the theory closer to the reality of students.

However, the approach to RT remains uncommon in Brazilian and global classrooms, as noted by Kaur et al. (2017); Pitts et al. (2013). A prevalent obstacle to the integration of RT into educational curricula is the inherent complexity of the subject, posing challenges for students in grasping its intricacies. Teachers

and students encounter various difficulties related to RT, with a primary concern revolving around the capacity to abstractly visualise relativistic situations (Ruggiero et al., 2021; Steier & Kersting, 2019; Yavaş & Kızılcık, 2016).

Given this educational landscape, an advantageous strategy for enhancing the learning process of RT could be the construction of mental representations by students, facilitating a deeper understanding of the theory and the associated phenomena.

Dual Coding Theory (DCT)

Einstein's Relativity Theory (RT) stands out for challenging various aspects of common sense, introducing novel notions of space and time that diverge from students' familiar experiences in their daily lives (Kraus, 2008). Consequently, the construction of mental representations may prove beneficial in enhancing students' grasp of the theory.

Mental representations are internal constructs individuals devise to capture their experiences. Human cognition, involving both nonverbal objects and language, operates through the simultaneous utilization of two coding systems: verbal, employing language, and nonverbal, utilizing imagery (Clark & Paivio, 1991).

> The structural representations of dual coding theory refer to relatively stable long-term memory information corresponding to perceptually identifiable objects and activities, both verbal and nonverbal (Paivio, 1986, p. 54).

These two autonomous systems are structurally and functionally distinct, enabling individual or combined activation. One system processes information and generates nonverbal representations, creating mental imagery, while the other system manages language and propositions. A proposition is defined as "an abstract subject-predicate statement that can be evaluated for its truth value within the bounds of a logical theoretical system" (Paivio, 1986, p. 9).

According to the DCT, mental representations originate from the internalization of external perceptual experiences and are stored in modality-specific representational structures as symbolic coding systems: the verbal one, represented by *logogens*, and the nonverbal one, represented by *imagens*. In verbal structures, smaller units progressively organize into larger ones. Conversely, nonverbal units are available simultaneously, facilitating parallel information processing.

Furthermore, "related context effects influence not only the relative activation of verbal and nonverbal systems but also the patterns of activation of verbal and nonverbal systems" (Clark & Paivio, 1991, p. 156). Therefore, the origin

and nature of external stimuli can impact the likelihood of which type of mental representation is employed. The activation of nonverbal representations correlates with the concreteness of stimuli and is linked to sensory experiences, meaning that words associated with concrete and sensory experiences tend to evoke mental imagery. Conversely, more abstract concepts, not easily linked to sensory experiences, tend to generate verbal representations, potentially in the form of propositions.

Then, both systems are important for an individual's cognition:

> Performance is mediated by the joint activity of verbal and nonverbal systems, with the relative contribution of each system depending on characteristics of the task and cognitive abilities and habits of the performer (Paivio, 1986, p. 201).

In summary, for a comprehensive understanding of a phenomenon, individuals use both information coding systems. Consequently, in our analysis, we aimed to identify the presence of mental representations – verbal (logically formed sentences) and nonverbal (imagery or mental simulations) – in students' cognitive structures following their participation in the activities.

Research Question

Given the aforementioned context, this research endeavours to identify whether and which mental representations are constructed and employed by high school students in the learning process of both Special and General Relativity Theory. Stemming from the research problem, "*What kind of mental representations do students use to process relativistic phenomena?*", we aimed to explore the coding systems utilized by students in grappling with relativistic phenomena related to both time and space.

Methodology

This research comprised two phases: initially, an exploration focused solely on Special Relativity (SR), followed by a subsequent phase incorporating General Relativity (GR). In both segments, high school senior students actively participated in the data collection, engaging with various instructional resources, including videos, animated gifs, mock-ups, discussions, and computer simulations. The activities encompassed diverse forms of representation to elucidate phenomena related to time and space, incorporating both verbal and nonverbal stimuli for the comprehensive study of all phenomena.

The first phase of the study took place during physics curricular classes in the second half of 2019. Students underwent pre-test and post-test questionnaires, leading to the selection of 14 students for individual interviews conducted by the teacher-researcher. The interviews followed the Repot Aloud protocol (Trevisan

et al., 2019), fostering a continuous dialogue between the interviewer and interviewee. During these interviews, students articulated their thoughts while recounting their experiences with the activities.

The second phase of the research involved eight students in an extracurricular short course during the second half of 2022. Throughout the course, students collaborated in small groups to perform various activities, stimulating discussions among them. The analysis of these activities included the examination of worksheets submitted by students, complemented by video recordings capturing their interactions. Utilizing this data, the students' dialogues were scrutinized to discern their reasoning processes.

Both the interviews in the first part and the recorded activities in the second part were fully transcribed. The transcripts underwent textual analysis to identify the verbal representations used by the students, taking the form of logically connected verbal sentences. This analysis drew upon characteristics of verbal representations and their usage as presented by Clark and Paivio (1991). Indicators of the presence of verbal representations included the use of extensive language to describe a situation, incorporation of verbal details, or auditory processing, suggesting the student was 'hearing' the information in their mind.

Verbal representations could manifest as complete sentences or even as single words, depending on the complexity of the situation. Additionally, "verbal representations are generally processed in a serial or sequential manner" (Clark & Paivio, 1991, p. 151). Hence, when employing verbal representations, students sequentially provide a verbal description of a situation, presenting a logical sequence of interconnected events.

The Depictive Gestural Analysis (Monaghan & Clement, 1999; Stephens & Clement, 2010) was also employed to explore the connection between descriptive gestures performed by students and their mental images, i.e., nonverbal representations. This analysis assumes a correlation between the depictive gestures executed by students and the mental imagery processes in their cognition.

In an effort to identify these depictive gestures, indicators such as students' hand movements during dialogue were used. Consequently, gestures serve as a means for students to externalize their thoughts, indicating their mental imagery or nonverbal representations. Depictive gestures can be seen as a parallel 'speech' that occurs alongside verbal communication. When jointly analysed, it becomes possible to discern the mental representations employed during the activities.

Results and Discussion

Through the analysis of results, it was possible to identify the use of both types

of mental representations, verbal and nonverbal, by students for both SR and GR. The results presented below show evidence of these mental representations first for SR and, afterwards, for GR.

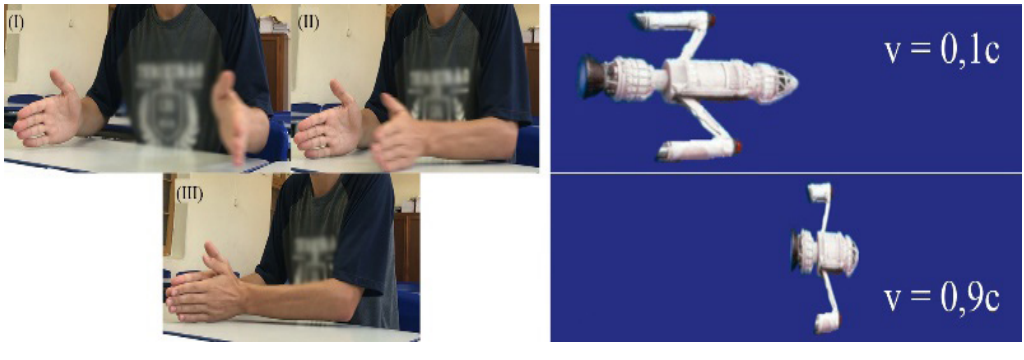
Special Relativity

Among the students who participated in the first part of the research, student A18 exhibited evidence of the use of nonverbal representations when grappling with the space contraction phenomenon. In addressing questions 1 and 3 of the tests, which pertained to the measured distance between two trucks in situations of varying relative velocities, A18 demonstrated an understanding of the relationship of this phenomenon with the speed of the reference frame.

> A18: . . . That the speed is high, so the higher the speed, the shorter the distance will be.
>
> E: . . . And do you imagine that?
>
> A18: Yes, I imagine [# NAV4 20:09] the path literally decreasing you know, as if it were in the GIF [*used in the slides' presentation*], but then the path and not the spaceship, right.

This student manifested a distinct mental image, specifically, a nonverbal representation constructed to comprehend the space contraction phenomenon. A18 described this representation as a spaceship decreasing in length as it moves. The ship, as conveyed by A18, along with his depictive gestures (Monaghan & Clement, 1999), bore a striking resemblance to an animated GIF used in slides, suggesting the existence of a mental simulation for this relativistic situation.

Figure 1. Gesture #NAV4 performed by A18 (left) where he places both hands in front of himself and moves the left hand close to the right hand, indicating the change in the spaceship length; and GIF from the slides (right).



Therefore, A18 translated the previously observed situation in class to the questions of the post-test ("the faster the truck was, the shorter its path would

be"), providing correct responses. Even though multiple representations were used during the classes to approach the space contraction, A18 primarily relied on the visual representation presented through the GIF to construct his personal mental representation of the phenomenon.

This student could construct a nonverbal representation, employing a mental simulation accessible to him offline to solve a distinct problem. In accordance with the DCT, concrete concepts like space, tied to sensory experiences, are more likely to evoke the use of the nonverbal system and, consequently, nonverbal representations.

Another participant from the first part of the research, student A14, similarly exhibited indications of using mental representations. A14 demonstrated an understanding of time dilation and employed verbal representations when elucidating her responses to questions 5 and 7 of the tests, which pertained to the time interval of a ship trip for two friends in situations with varying relative velocities.

> A14: I think it would be very bigger, like, more than one year of difference . . . from each other. . . I think about the little rocket, about the brother [*twins' paradox*], <u>because the brother who had traveled was younger than the brother who has stayed on Earth</u>.

It is possible to note that this student did not used images but relied on propositions to address time dilation. According to A14, she recalled a situation discussed during the activities, indicating the use of verbal representations. Additionally, A14 did not perform any depictive gestures nor allude to imagining anything directly, signifying the absence of mental images and nonverbal representations.

When questioned about what student recalled while answering the question, A14 answered 'about the brothers, because the brother that had gone traveling was younger than the brother that had stayed on Earth'. This suggests that the student verbalized the example and constructed a mental representation of the concept without relying on visual aids or simulations.

In the post-test, A14 accurately predicted the answers to questions concerning time dilation and articulated logical arguments during the interview, employing verbal reasoning to resolve the problem. The presence of a sequential and organized verbal structure suggests of the existence of a verbal mental representation (Clark & Paivio, 1991) of time dilation – where less time passes for a moving reference – indicating a sound understanding of the concept.

According to the DCT, this outcome aligns with expectations. Abstract concepts, such as time, which are not closely tied to sensory experiences, are likely to prompt the use of the verbal system and verbal representations.

General Relativity

Regarding the GR, the results obtained by the second part of the research mirrored those of the first. When addressing proposed question about spacetime distortion caused by the massive star Sirius, a group of three students discussed about the light bend around it. Evidence of the use of nonverbal representations by these students was identified as they engaged in collaborative depictive gestures during the discussion.

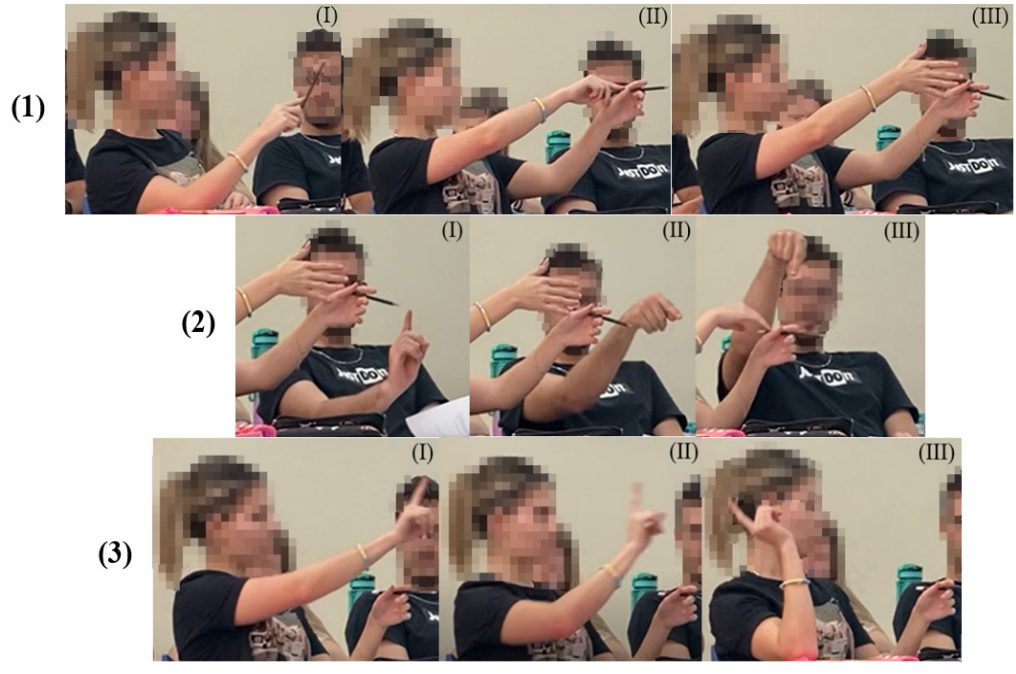
> P3: Yes, we saw this in another class, do you remember that she [*teacher*] explained that depending on the angle that you look, you can't see [#LUZ1 01:54] the light.
>
> P6: Yes, if I'm not wrong, you were seeing here [#LUZ2 01:58], and the light was here.
>
> P3: And then, it made a curve [#LUZ3 02:02].

In this activity, P3 and P6 were elucidating the concept of light bend to a third student, P10, who faced challenges in imagining the situation (Steier & Kersting, 2019). Throughout the dialogue, it became apparent that the students attempted to visualize the situation, employing collaborative depictive gestures to articulate their mental images, namely, their nonverbal mental representations. Moreover, the students referenced the classes where an animated GIF was used to illustrate this particular situation.

Despite the variety of representations employed to approach the light bending during the classes, the students primarily relied on the visual resource provided by the animated GIF to understand and explain the proposed situation. As the students described during the activity, and as suggested by their depictive gestures (Monaghan & Clement, 1999), they constructed a nonverbal mental representation dealing with the light bending phenomenon. Both P3 and P16 used their gestures to describe their imaginative processes and stablish a communication with P10.

The situation resented in the exercise that the students were solving, encompasses highly imagistic concepts, fostering the use of nonverbal representations (Clark & Paivio, 1991). Additionally, it is crucial to emphasize the collaborative process of imagination established by them (Steier & Kersting, 2019). Through this communication, the students were able to reach a collective conclusion regarding the solution of the problem.

Figure 2. (1) Gesture #LUZ1 performed by P3 pointing indicating the massive object, and with her hand indicating where an observer sees a light bean passing by it; (2) Gesture #LUZ2 performed by P16, interacting with P3 gesture, where he points where the light is emitted and where it would be observed; (3) Gesture #LUZ3 performed by P3 where she indicates the light bean trajectory.

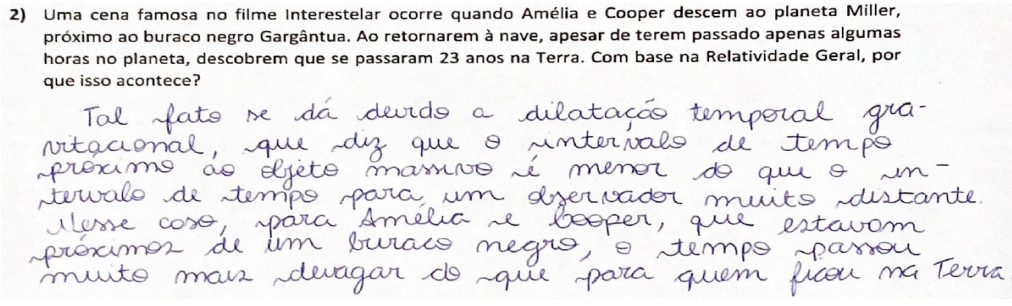


In responding to another question, which dealt with gravitational time dilation, student P9 exhibited some evidence of the use of verbal mental representations. The question involved a scene from the movie 'Interstellar', which had been discusses in the classes, and inquired about the difference in the time intervals between those who approached the black hole and those who remained on the spaceship.

While answering this question, no depictive gestures were identified during the students' dialogue. Furthermore, although students were encouraged to draw or visually present their answers using diagrams, P9 opted to provide her answer solely through text, describing the phenomenon that occurred.

In student's response, she explained that "the time interval near a massive object is smaller than the time interval for an observer far away". Notably, in P9's answer, she established a connection between the passage of time and the distance from a massive body. The interlinked sentences reflect a logical relationship among them.

Figure 3. P9's answer to the question about gravitational time dilation.


2) Uma cena famosa no filme Interestelar ocorre quando Amélia e Cooper descem ao planeta Miller, próximo ao buraco negro Gargântua. Ao retornarem à nave, apesar de terem passado apenas algumas horas no planeta, descobrem que se passaram 23 anos na Terra. Com base na Relatividade Geral, por que isso acontece?

Tal fato se dá devido a dilatação temporal gravitacional, que diz que o intervalo de tempo próximo ao objeto massivo é menor do que o intervalo de tempo para um observador muito distante. Nesse caso, para Amélia e Cooper, que estavam próximos de um buraco negro, o tempo passou muito mais devagar do que para quem ficou na Terra.

Hence, these indications suggest that this student constructed and employed verbal mental representations to address gravitational time dilation and comprehend it. This outcome aligns with expectations, given that dime dilation involves abstract concepts (such as time) not directly linked to sensory experiences. According to DCT, such concepts tend to favor the use of the verbal coding system (Paivio, 1986).

Conclusion

In this work, we emphasize the importance of incorporating Relativity Theory into school curricula. Given its complexity and the scarcity of observable consequences in everyday life, students often encounter significant difficulty in imagine relativistic phenomena (Ruggiero et al., 2021; Steier & Kersting, 2019). Therefore, we posited the hypothesis that the development of mental representations by students could enhance their comprehension of Relativity Theory. This study aims to identify the types of mental representations that emerge in students' learning processes.

The analysis of the results revealed the usage of both verbal (propositional) and non-verbal (imagistic) mental representations by students when dealing with SR and GR. It was also possible to note that verbal mental representations allowed the students to interpret more abstract phenomena involving time dilation for both Special and General Relativity. As time is an abstract concept not directly linked to sensory experiences, this outcome aligns with the Dual Coding Theory (Clark & Paivio, 1991; Paivio, 1986).

Conversely, nonverbal representations were associated with more concrete situations, primarily addressing space-related concepts, such as space contraction in SR and the light bending in GR. These mental representations manifested as mental imagery and mental simulations, aiding students in visualizing phenomena related to spatial concepts. This result is also consistent with the DCT, predicting that concrete phenomena, strongly linked to sensory experiences, tend to activate the nonverbal coding system.

It is crucial to observe that students who demonstrated an understanding of the

theory exhibited evidence of these mental representations. Then, it appears that both coding systems, verbal and nonverbal, are essential for comprehending Relativity Theory. Therefore, in response to the research question proposed, it is observed that students used both verbal (propositional) and nonverbal (imagistic) mental representations to process the relativistic phenomena approached, primarily involving time and space.

These findings underscore the importance of comprehending and analysing the mental representations students develop in their learning processes of complex theories. This information becomes pivotal for devising effective teaching strategies that foster a deeper understanding of the phenomena explained by these theories.

Furthermore, the analytical framework applied in this study can be extended to explore other themes within Modern and Contemporary Physics, as well as across various Science topics. Such an extension contributes to a better understanding of students' cognitive processes and facilitates the promotion of meaningful and hight-quality education.

Acknowledgement

This work was supported by the Coordination for the Improvement of Higher Education Personnel (CAPES), under grant 88881.846299/2023-01.

References

Boublil, S., Blair, D., & Treagust, D. F. (2023). Einstein's "happiest thought": Free-fall experiments enhance student learning of Einsteinian gravity in middle school. [Journal Article]. *Teaching Science*, *69*(1), 41-52. https://search.informit.org/doi/10.3316/informit.935555818844878

Clark, J. M., & Paivio, A. (1991). Dual Coding Theory and education. *Educational Psychology Review*, *3*(3), 62. https://doi.org/1040-726X/91/0900-0149506.50/0

Hartle, J. B. (2006). General relativity in the undergraduate physics curriculum. *American Journal of Physics*, *74*(1), 14-21. https://doi.org/10.1119/1.2110581

Henriksen, E. K., Bungum, B., Angell, C., Tellefsen, C. W., Fragat, T., & Boe, M. V. (2014). Relativity, quantum physics and philosophy in the upper secondary curriculum: challenges, opportunities and proposed approaches. *Physics Education*, *49*(6), 8. https://doi.org/10.1088/0031-9120/49

Kaur, T., Blair, D., Moschilla, J., Stannard, W., & Zadnik, M. (2017). Teaching Einsteinian physics at schools: part 1, models and analogies for relativity. *Physics Education*, *52*(6), 065012. https://doi.org/10.1088/1361-6552/aa83e4

Kersting, M. (2019). Free fall in curved spacetime—how to visualise gravity in General Relativity. *Physics Education*, *54*(3), 035008. https://doi.org/10.1088/1361-6552/ab08f5

Kraus, U. (2008). First-person visualizations of the special and General Theory of Relativity. *European Journal of Physics*, *29*(1), 1. https://doi.org/10.1088/0143-0807/29/1/001

Monaghan, J. M., & Clement, J. (1999). Use of a computer simulation to develop mental simulations for understanding relative motion concepts. *International Journal of Science Education*, *21*(9), 921-944. https://doi.org/10.1080/095006999290237

Paivio, A. (1986). *Mental Representations: a dual coding approach*. Oxford University Press.

Pitts, M., Venville, G., Blair, D., & Zadnik, M. G. (2013). An exploratory study to investigate the impact of an enrichment program on aspects of Einsteinian Physics on Year 6 students. *Research in Science Education*, *44*, 363 - 388.

Ruggiero, M. L., Mattiello, S., & Leone, M. (2021). Physics for the masses: teaching Einsteinian gravity in primary school. *Physics Education*, *56*.

Steier, R., & Kersting, M. (2019). Metaimagining and embodied conceptions of spacetime. *Cognition and Instruction*, *37*(2), 145-168. https://doi.org/10.1080/07370008.2019.1580711

Stephens, A. L., & Clement, J. J. (2010). Documenting the use of expert scientific reasoning processes by high school physics students. *Physical Review Special Topics - Physics Education Research*, *6*(2), 020122. https://doi.org/10.1103/PhysRevSTPER.6.020122

Trevisan, R., Serrano, A., Wolff, J., & Ramos, A. (2019). Peeking into students' mental imagery: the Report Aloud technique in science education research. *Ciência & Educação*, *25*(3), 647-664. https://doi.org/10.1590/1516-731320190030004

Yavaş, P. Ü., & Kızılcık, H. Ş. (2016). Pre-service physics teachers' difficulties in understanding Special Relativity topics. *European Journal of Physics Education*, *7*, 13-24.

A Human Orrery to Help 11 Years Old Students Investigate The Shape of Planetary Orbits

Emmanuel Rollinde, Clément Maisch

[1]Cergy Paris Université, Université de Paris, France

We present a learning session designed around the use of a Human Orrery, a realistic representation of the solar system. Moving on the Human Orrery, students play the role of celestial objects such as the internal planets and a comet. By moving along the orbits, 11 years old students should notice and then explain the non-circular attribute of the orbits of planets. The Didactics Engineering has been mobilized to design the session following several didactical variables such as: the size of the Orrery, the eccentricities of the orbits or the proposed instruments. During the session, students make hypothesis, define protocols to prove their hypothesis and discuss the measures they obtain to confirm or discard their hypothesis. In the course of this modelling process, the circulation of their conceptions of a circle have been analysed. The observations of three classrooms involved in this session indicate that the students change from a continuous image of the circle, related to visual estimation and drawing with a pair of compasses, to a discrete understanding related to the definition of a circle as the set of points at equal distance from a centre. This session, led by two researchers of physics education for the research presented here, has also been implemented by science teachers with the same efficiency and conclusions.

Keywords: geometry, interdisciplinarity, astronomy

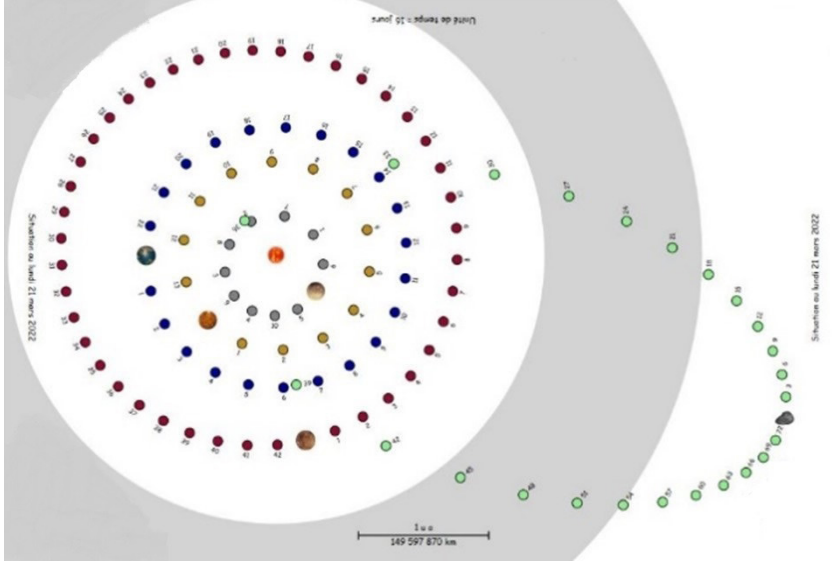
Introduction

We investigate the use of a Human Orrery (HO, Figure 1), to design a pedagogical session that combines geometrical activities and measurements in a modelling process (Sensevy, 2008; Yvain-Prébiski, 2021). This representation of the solar system in the HO is a realistic[2] illustration of the movement of celestial objects – the four internal planets and a comet in our setting. Such representations have been vastly used in the literature of astronomy education (e.g.: Asher at al., 2007; Gould, 2005; Newbury, 2010). Our approach is specific in three key elements: (i) the focus of the learning session is on the science and mathematics knowledge that are embedded in the description of the Solar System (such as kinematics in Rollinde, 2017 and Rollinde et al. 2021; proportionality in Abboud

2 Positions of the objects at different dates are reported using accurate astronomical data

et al., 2019; or geometry in Nechache & Rollinde, 2023 and Rollinde & Maisch, 2023 – both in French - and this paper). rather than on the astronomy (ii) The design of the learning session is based on the embodied design in mathematics and science education (Abrahamson et al. 2020; Kersting et al. 2023) with an explicit connection between the pieces of knowledge and the physical actions of the learner. (iii) Usually, HO are built and used with schools in the premises of observatories or museum. We have then developed two other designs for the HO. One is a printed version on a A3 paper (printed Orrery, PO), and a second one is printed on a tarp with a scale of one metre for one astronomical unit (the distance from the Sun to the Earth), fitting with the definition of a Human Orrery (HO). Those two designs ease the dissemination of learning sessions in schools and even inside a classroom. Besides the actions with a PO or on a tarp are complementary (see Rollinde et al., 2021). On Human Orrery (HO), students walk on it so that they are able to perceive the movements of celestial bodies along their orbits. On the Printed Orrery (PO), students used coins or their fingers to move along the orbits. They have a global point of view of the orbits at once. Due to the size of the HO, the mathematical instruments (ruler, compass) are not the one used in the standard activity in a classroom which can be used with the PO.

Figure 1. Successive positions of celestial objects are displayed with discs placed every 16 days along each orbit.



The focus of this paper is on the circle, and its properties in relation with the curriculum of a junior high school (11 years old students). The description of planetary orbits is an ideal subject to combine geometrical models and experimental methods of modelling. This is the case from a historical point of

view (Linton, 2004). We demonstrate here that is can also be used in a teaching session for 11 years old students. Through a modelling process based on the use of the HO and the PO, we will analyse the circulation of students' knowledge and conceptions of a circle.

Method

The session was designed following the Didactics Engineering method (Artigue, 2014) together with the principles of embodied design. The knowledge is defined first. Situations are conceived with different didactical variable (size of Orrery, eccentricity of orbits and available instruments) to allow a conceptual change in student's understanding of knowledge (circles here).

Embodied design

The question of the body from the point of view of learning is the subject of a field of research initiated in the 2000s in mathematics education (Abrahamson et al., 2020) and then science education (Kersting et al., 2023), based on the paradigm of "embodied cognition" (Varela et al., 1991). According to this paradigm, the individual's interactions with his or her environment are constitutive of cognition; abstract concepts are based on these sensory-motor experiences. Here, on the Human Orrery, students are engaged in a collective activity, directed towards the same object. They will enact planets through a multimodal semiotic activity combining language and sensory-motor actions, interacting with cultural artefacts to construct scientific concepts and enable the effective production of meaning (Arzarello & Robutti, 2010). During a session, each student brings its own individual pre-scholar model of the Solar System and create a singular cognitive model through interaction with the environment (material and social). This should allow the emergence of a collective model. The session is managed so that activities of learners and of the group of learners lead to a collective model that is consistent with the precursor scholar model as described in the next section (namely, non-circular orbit or, *a minima,* orbits that are not centred on the Sun).

Models of circles in astronomy education

The first law of Kepler tells that "All orbits draw an ellipse with the Sun as one of the foci". Their eccentricity depends on the object but is never equals to zero. Hence, the Sun is never the exact mathematical centre of the orbits. The notion of ellipse is not present in the French curricula before the University. Hence, our focus is rather on the figure of the circle. According to the literature in mathematics education (from Artigue, 1982 to Bulf et al., 2021), three main conceptions about the circle may be expected to be used by students after primary school. (C1) A circle is the geometrical figure drawn with a pair of compasses

(we use hereafter "compass" only). This conception has to be connected with the definition (D1) of a circle as the set of discrete points placed at equal distance of a centre. (C2) A circle is a continuous, round, line, which must be connected with the definition (D2) of a closed curve with a constant radius. Last, (C3) a circle has two parts which could be doubled exactly. This will be extended into the definition (D3) of a circle as a figure with an infinity of axis of symmetry.

The lesson is then designed in accordance with the modelling process defined by Schwarz et al. (2009) or Nielsen and Nielsen (2021) with a dialogue between the empiric world (a place to perceive and to measure) and the models (a place full of quantities and relations). The practice of modelling process involves (among others) creation of models, using models through prediction, evaluation and revision of models… To prove that orbits are not circular, one may search for the centre through the use of three bisectors. Yet, students at junior high school are unlikely to know this procedure. Hence, we assume they will focus on operations dealing with measurement of distances. To prove first that the Sun is not the centre of the orbit, one measure distances from different points along the orbit to the Sun. Uncertainties are then related to the size of each disc (the centre is not identified) and to the units of the ruler. We estimate uncertainties to be of the order of 6-7 mm on the Human Orrery and 2-3 mm on the Printed Orrery. The difference between the largest and smallest distance to the Sun for each orbit is related to its characteristics (major and minor axis, eccentricity of the ellipse) as given in Table 1. It appears that the difference is larger than the uncertainty for all orbits using the Human Orrery, but only for Venus and Earth on the Printed Orrery. Hence, in our situation, students will first test their hypothesis on the Printed Orrery, and may be convinced the Sun is the centre. They will then proceed to the Human Orrery which will contradict their findings.

Once it has been proven that the Sun is not the centre on the Human Orrery, one may wonder if the orbit is a circle but with a different centre. Since the Orrery does not provide a continuous line but a series of discs, one may not find diameter through the maximum length of a chord. Yet, through a trial-and-error procedure, students may search for a point that is at equal distance to each disc. Table 1 (last line) indicate the difference in distance to the true centre of the ellipse (which is the minimum difference that one may get). It appears that – besides the orbit of the comet Encke – only the orbit of Mercury can be proven not to be a circle. Other orbits have a too small eccentricity to differ significantly from a circle.

The conclusion must then be that "the orbit of Mercury is not circular, while the orbits of Venus, Earth and Mars are consistent with a circle whose centre is not the Sun".

Table 1: Table of characteristics of each orbit (major and minor axis, eccentricity), and difference between the largest and smallest distance to the Sun.

		Mercury	Venus	Earth	Mars	Encke
Semi major axis, a/2	[UA]	0,387	0,72332	1,0000	1,523	2,218
Semi minor axis, b/2	[UA]	0,379	0,72330	0,9998	1,517	1,178
Eccentricity, e		0,205	0,007	0,016	0,093	0,847
Distance difference to the Sun (Δl_{orbit} =2ae)	[UA]	0,158	0,010	0,032	0,285	3,757
	Human Orrery [cm]	15,8	1,0	3,2	28,5	375,7
	Printed Orrery [cm]	0,9	0,0*	0,2*	1,6	21,4
Distance difference to the centre of the ellipse (Δl_{orbit} =a-b)	[UA]	0,008	0,00002	0,0002	0,007	1,039
	Human Orrery [cm]	0,8	0,0*	0,0*	0,7*	103,9
	Printed Orrery [cm]	0,0*	0,0*	0,0*	0,0*	5,9

Each difference in distance is given in Astronomical Units and the conversion on the scale of the human Orrery and of the Printed Orrery. Differences in distance on the Orrery that are lower than our estimate of the measurement uncertainty are indicated by an asterisk.

Design of the session

Thirty-eight 11-years old students (three classes of the French “6ème” at junior high school) followed an experimental science classroom, led by two of the authors who are physics education researchers. Students are guided into a modelling process, grounded on the description above, during which their conceptions of the circle may be challenged. Firstly, we investigate how students relate their kinaesthetic perceptions as they move and their visual perceptions as they see the orbits to the properties of the circle. Secondly, how they confirm the predictions related to those properties through an act of measurement. Thirdly, how they modify their initial model accounting for the results of the measurement at bodily size environment (HO).

In phase 1, orbits of planet on the HO are experimented by moving on it without

measurement devices. Each student is asked to make its own hypothesis about their shape, and write it on a paper form. In phase 2, each student proposes a protocol to test its hypothesis. Any measurement device may be used, such as a compass, a ruler and a string. This protocol is tried out with a series of six squared pictures, about 10 cm-large, displaying discs placed along imaginary elliptical orbits as in the HO. Only one of those orbits is a circle. Students then discuss in groups of 4 about their protocol, write it down on the form and then apply it individually on an Orrery printed in a A4 size. Afterward they discuss by groups if their hypothesis is confirmed or discarded. In phase 3, all groups go together on the tarp (HO) and are asked to test their protocol there and to write down their new conclusions if they changed.

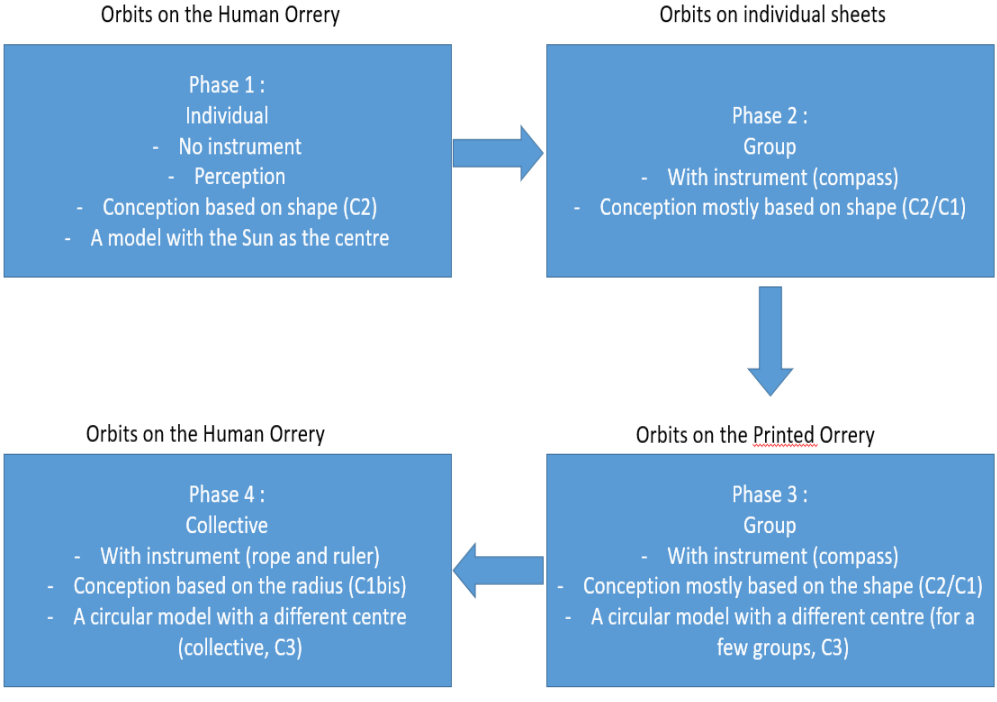
Collection of the data

All along the session, students must complete the forms individually, make measurements and discuss about their hypothesis about circularity of the orbits. We collected the 38 forms filled by the students and we videotaped their moves on the tarp (phases 1 and 3) as well as the table work of one group during phase 2. Answers and explanations (oral and written) as well as the choice of the measurement devices are used to determine which conceptions emerge in each phase. We looked at specific gestures such as for C1: showing a series of discs the way they draw circles or link discs on the form or in the PO or HO (e.g., with a rope) or speaking of a round instead of a circle; for C2 by taking in account distance between the Sun and a disc, using a compass or a rope to compare the distances between discs and the Sun. We also investigate how students consider errors in measurements to make a decision about their hypothesis.

Results

Hereafter, the global attitude that was demonstrated by almost all but a few students is described. We first investigate whether the modelling path followed by the students through the three phases was successfully implemented in the situation. The evolution of the conceptions is clearly revealed by the activities of the students (Figure 2). In Phase 1, they hypothesized a circular or “round” shape for planetary orbits and an unknown, not round, shape for the orbit of the comet. As expected, this hypothesis is grounded on their visual perception (“it looks like a round”), their kinaesthetic perception as they enact the planets (“planets turn around the Sun”). We may also expect that the circular shape is favoured by the scholar model taught in primary school.

Figure 2 - Description of the four phases of the Lesson and the evolution of the conceptions about the circle



Their protocols to investigate if orbits are indeed circles were then based on the use of a compass. Although they did not mention the Sun as being the centre of planetary orbits in their hypothesis, we observe that they spontaneously place the "needle point" of their compass in the estimated centre of the Sun. The belief that the Sun is the centre of the orbits – or more probably, "the centre of the Solar System as a whole" understood as the centre of each single orbit, is then an implicit. They discussed then if the "pencil lead" of the compass goes through all discs. The successive position of planets along their orbit should be identified as the centre of each individual disc (Figure 1). However, only a few students discuss if the circle drawn by the compass intersects the different discs at the same position. Given the accuracy of their procedures, orbits of Earth and Venus could not be distinguished from a circle centred on the Sun (one circle may indeed intersect all discs in "nearly-identical" points). On the contrary, the drawing of a circle centred on the Sun could not go through all points along the orbits of Mercury or Mars. As they realised this, some students were surprised, but all accepted to discard their initial hypothesis. Only two groups (among 10) tried to place their compass on another point to adjust the orbits with a circle whose centre would not be the Sun (which is the right way to proceed). As they proceed into the HO on a large tarp (phase 3), they could not use their compass

due to the size of the tarp. Some tried to place a rope along all points belonging to one orbit, but could not make use of this procedure further. Then, all classes ended up collectively by using a rope or a ruler to evaluate the distance from the Sun to each point along one orbit. Few students in each class were then able to discuss uncertainties related to the choice made as the two ends of the rope are set at the centre of the Sun for one and of each disc for the other. Yet, they all realized that Venus' orbit was still consistent with a circle centred on the Sun while the orbit of Earth was a subject of debate.

Along the modelling process, the circulation of students' conceptions (Figure 2) could be clearly observed. Results are again very homogeneous for all students. The circle is conceived first as a continuous and round line, observed visually or while moving along it (phase 1, C2). It is then linked to the drawing with a compass (phase 2, C1). The large size of the HO (phase 3) forces the students to connect the continuous path of the compass (C1 and C2) with the definition (D1) based on the radius. We note that the conception C3 did not appear during the session.

Conclusion and Discussion

Through the analysis of the implemented session, we observed how students have followed a modelling process, while dealing with geometrical conceptions of the circle. According to Yvain-Prebiski (2021) modelling requires starting from an extra mathematical problem, building a model, making it work and being able to compare its results with the modelled situation. Planetary orbits are simple but visually appealing shapes. The shapes of planetary orbits can be modelled using simple mathematical figures that reflect students' initial conceptions. The mathematical properties of these shapes are accessible to 9 to 11 years old pupils, making them aware of the need to make choices and that, depending on the choices made, the parameters of the model chosen (choice of the centre) will not be the same ("horizontal" mathematisation process, ibid) and the tools used to test the prediction ("vertical" mathematisation process) will not be the same either (compass or ruler). In our case, students are guided towards the circular model by the situation, their knowledge of geometry and the images of orbits they have encountered in the past. On the other hand, we can see that they quite quickly have the choice of enriching this model by not imposing the position of the centre of the circle. The mathematical treatment of the model can be simple if the centre is known (using a compass or a string) or more complex if the centre is not known. The essential point here is that the pupils have several (two) mathematical models at their disposal which they know well enough to make predictions about the results of their measurement. We consider that the use of different planetary orbits and of the alternation from Orreries of different sizes has allowed this practice of modelling to occur.

The same session was implemented one year later by science teachers in the same school, and led to the same conclusions. As expected, the pedagogical focus of science teachers was on measurement activity, and modelling process. We advocate that such teaching sessions using the Human Orrery would be more adequately implemented in the future as collaboration projects involving both science and mathematics teachers.

Acknowledgement

This work was funding by the ARISTARCHUS Erasmus+ project (http://aristarchusproject.eu) and the CY Cergy Paris Université Emergence project (E-ESMEA).

References

Abboud, M. & Rollinde, E. (2021) Les mathématiques du Système Solaire en plein air. Le planétaire humain au collège. *Repères IREM, IREM*, 27-52. ⟨HAL-03293334⟩.

Abboud, M., Hoppenot, P., & Rollinde, E. (2019). *Enhancing mathematics and science learning through the use of a Human Orrery*. CERME11, Feb 2019, Utrecht, Netherlands. ⟨hal-02303967⟩

Abrahamson, D., Nathan, M. J., Williams-Pierce, C., Walkington, C., Ottmar, E. R., Soto, H., & Alibali, M. W. (2020, August*).* The future of embodied design for mathematics teaching and learning. In *Frontiers in Education* (Vol. 5, p. 147). Frontiers Media SA.

Artigue, M. (1982). À propos des conceptions du cercle, *Grand N, 27,* 45-72.

Artigue, M. (2014). Didactic engineering in mathematics education*.* In S. Lerman (Ed.), *Encyclopedia of mathematics education*. Springer.

Arzarello, F., & Robutti, O. (2010). Multimodality in multi-representational environments. *ZDM*, *42*, 715-731.

Bulf, C., Celi, V., Millon-Fauré, K., Beaugrand, C., & Dias, C. M. (2021). Tracé du cercle et circulation des discours (première partie) approche didactique des (inter) actions langagières et matérielles. *Petit x*, *114*, 3-37.

Kersting, M., Amin, T. G., Euler, E., Gregorcic, B., Haglund, J., Hardahl, L. K., & Steier, R. (2023). What is the role of the body in science education? A conversation between traditions. *Science & Education*, 1-40.

Linton, C. M. (2004). *From Eudoxus to Einstein: a history of mathematical astronomy*. Cambridge University Press.

Nielsen, S. S., & Nielsen, J. A. (2021). A competence-oriented approach to models and modelling in lower secondary science education: practices and rationales among Danish teachers. Research in Science Education, 51(Suppl 2), 565-593.Rollinde, E., (2017). Learning science through enacted astronomy*. International Journal of Science and mathematics education*, 17(2), 237-252. https://doi.org/10.1007/s10763-017-9865-8

Rollinde, E., Decamp, N., & Derniaux, C. (2021). Should frames of reference be enacted in astronomy instruction? *Physical Rewiew Physics Education Research*, 17(1), 013105. https://doi.org/10.1103/physrevphyseducres.17.013105

Rollinde, E., & Maisch, C. (2023). Les orbites planétaires sont-elles circulaires ? : Une ingénierie didactique associant mathématiques et sciences en cycle 3. *Grand N, Revue de mathématiques, de sciences et technologie pour les maîtres de l'enseignement primaire*, 111, pp.5-39. ⟨hal-04227613⟩

Schwarz, C., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., Shwartz, Y., Hug, B., & Krajcik, J. (2009). Developing a learning progression for scientific modeling : Making scientific modeling accessible and meaningful for learners. Journal of Research in Science Teaching, 46(6), 632654. https://doi.org/10.1002/tea.20311

Sensevy, G., Tiberghien, A., Santini, J., Laubé, S., & Griggs, P. (2008), An epistemological approach to modeling: Cases studies and implications for sciences teaching, *Sciences education*, *92*(3), 424–446.

Varela, F., Thompson, E., & Rosch, E. (1991). The embodied mind: cognitive science and human experience mit press. *Cambridge, Massachusetts*.

Yvain-Prébiski S. (2021). *Didactical adaptation of professional practice of modelling: a case study* In International Perspectives on the Teaching and Learning of Mathematical Modelling, Kaiser &al, Springer Book Series

The Literature Review on Teaching and Learning the Properties of Matter in Physics Education

Ayçin Ünal, Lorenzo Gianni Santi, Marisa Michelini[1], Mehmet Fatih Taşar[2]

[1]University of Udine, Italy

[2]Georgia State University, Atlanta, USA

The purpose of this literature review is to identify, list, and analyze published research studies on teaching and learning the topic of "properties of matter" in relation to physics education. The search was conducted on the Web of Science and Google Scholar databases. Fifteen publications have been identified that comply with the search criteria. Analysis of these articles yielded that published studies mostly employed qualitative research methods. The analysis further yielded four themes in which these articles can be grouped. As a result of this research, it helped to identify existing gaps in the literature and suggested further research on teaching and learning properties of matter.

Keywords: Literature review, physics education research, properties of matter

Introduction

One of the most critical roles in scientific literacy is played by the concept of matter that is one of all scientific concepts (Harrison, & Treagust, 2002). Making sensible choices in everyday life also needs a comprehension of the structure and properties of matter (Hadenfeldt et al., 2014). Since Piaget, the foundation of scientific education research has revolved around enhancing children's capacity to articulate the properties of matter and its transformations (Piaget, & Inhelder, 1974). Besides all these, when the curricula from various countries were examined, it was found that the properties of the matter were explained in a fragmented way (MEB, 2018a; MEB 2018b; MIUR, 2018; NGSS, n.d.). On the other hand, research in the field of teaching and learning properties of matter is abundant in the literature. But upon reviewing educational studies concerning the properties of matter, it becomes evident that they predominantly revolve around chemistry education, and mostly focus on the structure of matter. For instance, a study published in 2013 suggested a way to conceptualize high school chemistry education by developing student understanding about matter, energy, and models (Liu, 2013). Another research conducted in 2012 focused on improving the instruction and comprehension of the subject of matter and its

characteristics among students in the sixth grade (Pimthong, et al., 2012). In the literature, there are also studies that focus on the progress of student learning and uncover conceptual frameworks to address this progress (Krnel, Watson, & Glažar, 1998; Krnel, Watson, & Glažar, 2005). Andersson (1990) proposed two main categories for organizing the structure and changes of matter, with the aim of tracking learning progress. These categories include:

i. Conceptions related to the particle nature of matter encompassing ideas about atoms, molecules, and particle systems.

ii. Everyday conceptions regarding the matter and its alterations, encompassing concepts related to changes in physical states, and the conservation of matter, chemical reactions.

Hadenfeldt et al. (2016) and Liu, &Lesniak (2005) adopted Andersson's (1990) suggestion and further developed four conceptual dimensions for understanding matter, which are:

a) Structure and composition of matter.

b) Physical properties and changes of matter.

c) Chemical properties and changes of matter.

d) Conservation of matter.

Although there have been studies in the literature investigating the teaching of matter properties (Andersson, 1990; Hadenfeldt, 2016), in recent years, no study has found a literature review on this topic. Given its fundamental relevance in the realm of physics, it is essential to explore the most recent research, pedagogical methodologies, and innovative approaches that contribute to elevating student engagement, conceptual understanding, and practical application of physics principles. Due to the significance of this topic, we aimed to conduct a literature review to explore physics education studies that focus on the teaching and learning of the properties of matter. In this context, we addressed the following research questions:

1. What is the distribution of research on the properties of matter over the years?
2. What are the commonly employed methods in research related to the teaching and learning of the properties of matter?
3. What are the study groups or samples typically involved in physics education research on the properties of matter?
4. What are the key research issues and themes in physics education research on the properties of matter?

Method

We employed the method of a literature review in this investigation. The systematic literature review method can identify the literature gaps and what we need to know about that field (Petticrew& Roberts, 2006). The focus of the study should be decided before beginning a literature review and once relevant studies have been found, appropriate studies should be chosen using inclusion and exclusion criteria, and the articles should be assessed (Oxman, 1994).

Literature Search Strategy

The initial step in our literature review involved formulating a clear focus for the study. We defined the scope of our investigation to encompass the properties of matter, a fundamental concept widely studied across various disciplines. To ensure a comprehensive and rigorous review, we limited our search to include research conducted within the last 10 years, written in English, and published in peer-reviewed journals.

To gather relevant scholarly works, we conducted searches in the Web of Science® and ERIC databases using the keyword "properties of matter." Our initial search yielded 1133 results from Web of Science® and 287 results from ERIC. To refine our search and ensure relevance to the field of physics education, we further narrowed down our focus by selecting articles categorized under "physics education," "education & educational research," or "remote research education" within the Web of Science® categories. We reached 20 studies in this way in Web of Science®. Likewise, when we searched for the word "properties of matter" in ERIC and selected the studies done in the last 10 years, we found 37 articles.

Inclusion and Exclusion Criteria

The selection of appropriate studies was carried out through a meticulous process of inclusion and exclusion criteria. These criteria were established to ensure that the chosen studies align closely with the objectives of our review. Articles that were duplicates or not directly related to physics education were excluded from consideration. The process of selecting articles based on these criteria is visually outlined in Figure 1.

Final Article Selection

Upon applying the inclusion and exclusion criteria, we identified a final set of 15 articles from the Web of Science® and ERIC databases that met our review criteria. These selected articles constitute the core body of literature that we thoroughly analyzed and synthesized in our review.

Data Analysis

The data analysis process for this literature review on the properties of matter in physics education involved a systematic and rigorous approach to ensure the extraction of meaningful insights and reliable themes from the collected articles. By employing independent theme creation, collaborative discussions, and reliability checks, the analysis aimed to mitigate bias and enhance the credibility of the identified themes. The data analysis process for this literature review involved the following steps:

Independent Theme Creation: Upon selection, the articles were thoroughly read by two independent researchers. Each researcher independently identified and created initial themes based on the content of the articles. This process allowed for a comprehensive exploration of the diverse concepts, perspectives, and findings presented in the literature.

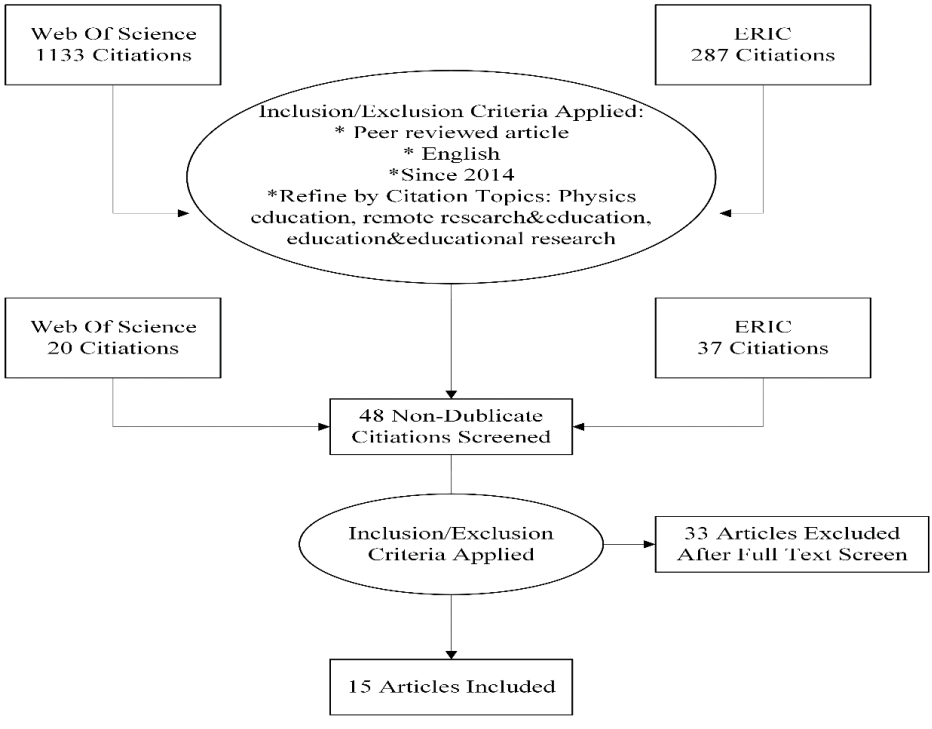
Code Generation: Following the creation of initial themes, the researchers proceeded to generate codes for each identified theme. Codes were specific keywords or short phrases that captured the essence of the content related to each theme within the articles. This step helped in systematically organizing the data for further analysis.

Theme Discussion and Consensus: In order to ensure reliability and consistency in the identification of themes, the two researchers engaged in a thorough discussion. They compared and contrasted the independently generated codes and themes, aiming to reach a consensus on the final set of themes that accurately represented the content of the articles.

Theme Determination: Through the collaborative discussion, the researchers refined and finalized the themes that best encapsulated the properties of matter in physics education as discussed in the articles. This iterative process of theme determination contributed to the establishment of a reliable framework for analysis.

Reliability Check: To enhance the reliability of the data analysis, a meeting was held specifically to review the identified themes and codes. This meeting focused on validating the accuracy and consistency of the themes. Adjustments were made where necessary to ensure a coherent and accurate representation of the data.

Figure 1. Inclusion and exclusion criteria



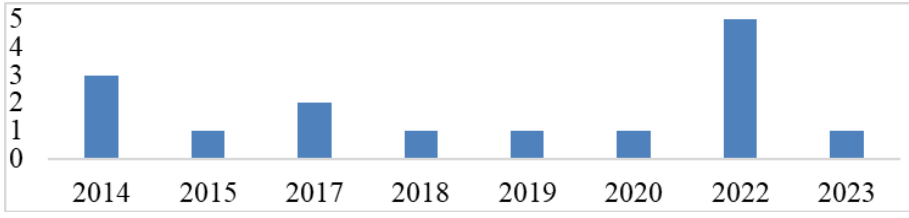
Results

This research on teaching and learning the properties of matter in the field of physics education involves a comprehensive review of 15 articles. These scientific studies have been meticulously examined and analyzed to enhance our understanding of this important topic of physics education. The findings from the research are presented in this section.

Years

Upon analyzing the distribution of studies across different years, it was observed that the majority of studies were conducted in the year 2022. This finding suggests a notable focus on research activities during that particular period about teaching and learning the properties of matter in physics education. Figure 3 visually represents the distribution of studies across years, providing a clear visualization of the trend.

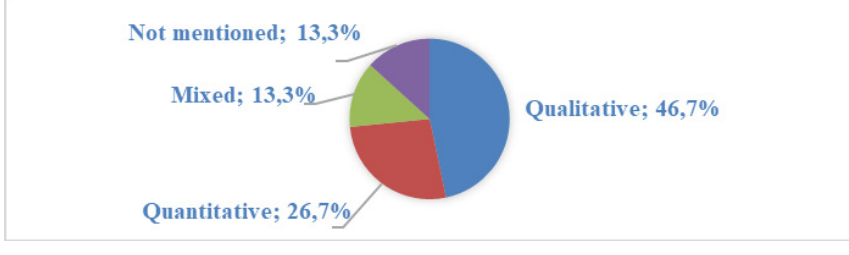
Figure 3. Distribution of articles across years



Research Methods

Upon examining the studies included in the analysis, it was observed that various research methods were employed. Out of the total number of studies, 7 (46,7%) utilized qualitative methods, 4 (26,7%) employed quantitative methods, and 2 (13,3%) employed mixed methods. However, there were also two studies (13,3%) in which the method used was not explicitly mentioned. The distribution of these methods can be visualized in Figure 4.

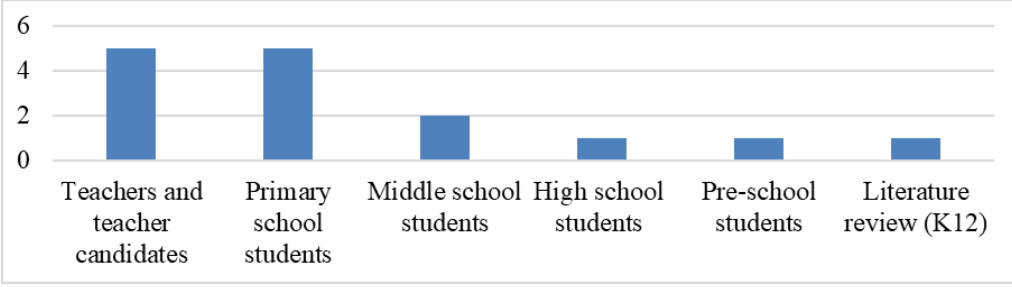
Figure 4. Distribution of research methods used in the studies



Sample/study groups

The studies that were reviewed revealed a wide variety of research focuses and targeted different study groups. Five studies were centered around teachers and teacher candidates, while five studies specifically targeted primary school students and another two studies focused on middle school students. Additionally, one study concentrated on high school students and one study stood out for its emphasis on pre-school students. One study took a unique approach and conducted a comprehensive literature review, offering a valuable resource for researchers and practitioners in the field. In summary, the reviewed studies encompassed a range of study groups, including teachers, teacher candidates, primary school students, middle school students, pre-school students, and high school students. The distribution of the studies according to the study/sample groups is given in Figure 5.

Figure 5. Distribution of articles across sample/study groups



Research topics/themes

We performed a content analysis by looking at the physics education research on the properties of matter. Potential codes that emerged from the investigations were identified by the two researchers who led the systematic review. We reviewed the significance of these codes and compared/verified the difference between the proposed codes at later meetings. Then, we identified the shared characteristics among the codes and developed the categories. When we categorized the studies according to the research issues and questions related to the teaching and learning properties of matter, four themes emerged.

Student ideas and conceptual changes: Findings from studies investigating students' conceptual understanding of the properties of matter and student ideas reveal the need for effective instructional strategies to promote conceptual change. These studies contribute to the development of instructional interventions aimed at enhancing students' understanding of matter properties.

One effective approach highlighted by Kara and Kingir (2022) is the model-based science writing heuristic approach. This approach improves students' conceptual understanding of matter properties and their ability to construct model-based argumentations. By engaging students in writing activities that involve constructing models and providing scientific explanations, this approach facilitates conceptual change.

Stamenkovski and Zajkov (2017) have shown that students have a poor understanding of the concept of mass and often confuse it with the concepts of density and weight. The article suggests the need for curriculum and textbook revisions, as well as improvements in the teaching of physics concepts, to address these conceptual confuses effectively.

Wilcox et al. (2022) conducted a study focusing on a three-day lesson about the properties of matter. In this research, they have emphasized the importance of teaching crosscutting concepts and the nature of science ideas through inquiry experiences. They found that hands-on activities and inquiry-based learning were effective in promoting conceptual change. According to this research result,

by actively participating in experiments and reflecting on their observations, students gained a deeper understanding of matter properties and developed a more accurate scientific perspective. To address conceptual shifts, practice-based assessments, model-based methods, and inquiry-based experiences can be utilized. Adams and Feagin (2017) conducted a study where they designed a lesson using the 5E method to enable them to explore scientific definitions of matter through both intensive and comprehensive properties. However, their investigation unveiled situations where students encountered difficulties in grasping the attributes of matter. They utilized both informal and formal evaluations to measure the students' learning progress in the research. The informal assessment encompassed group discussions and small-group tasks. The official assessment was executed via the Pirate Adventure app. The outcomes of the evaluation illustrated that, at the conclusion of the lesson incorporating the 5E method and hands-on exercises, students exhibited the ability to comprehend the properties of matter and apply this knowledge to classify objects.

Strategy and method in educational path: Within the properties of matter domain, studies under this theme focus on investigating teaching methods and strategies. In this researches, researchers have explored instructional strategies and approaches that facilitate students' understanding of the properties of matter, such as density, and states of matter and they have examined the impact of some activities, teaching methods and strategies on student engagement, conceptual understanding. The findings from these studies provide guidance for educators on selecting and implementing effective instructional practices in the teaching of properties of matter.

The research conducted by Allen and Rogers (2015) highlights the importance of the CER (Claim, Evidence, and Reasoning) framework was introduced as a tool for students to communicate their understanding through writing. In this study, an activity was designed and described, using the CER framework for upper elementary students to help them understand how the properties of matter affect light and sound, and their correlation. The authors argue that this approach can help students to develop a deeper understanding of scientific concepts and to improve their scientific literacy. Adams and Feagin (2017) have described a lesson that focuses on the intensive and extensive properties of matter and continues to reinforce these concepts throughout the school year. Kácovský (2019) has presented ten simple thermal imaging experiments designed to teach upper-secondary and younger students about the thermal properties of matter. Zorlu and Sezek (2020) have investigated the effects of combining cooperative learning and modeling-based teaching methods in teaching matter and thermal phenomena and the particle structure to improve students' academic achievement and science process skills.

Kara and Kingir (2022) have explored the use of the model-based science

writing heuristic approach to improve students' conceptual understanding and construction of model-based argumentations while teaching properties of matter and lighting and sound technologies units in elementary school science. The researchers discovered that the treatment group exhibited a notable edge compared to the control group in their performance on concept tests. This suggests a significant enhancement in their skill to build arguments utilizing models while applying the model-based science writing heuristic method.

Akkus and Doymus (2022) have compared the effectiveness of Reading Writing and Presentation (RWP) and Subject Jigsaw Method (JG) with the control group in terms of academic achievement about the properties of matter in 7th-grade students. After they implemented the study, they gave the students a posttest and found that both the RWP and JG groups performed statistically better than the control group.

Avcı (2022) examined the impact of the jigsaw technique, enriched with educational games, on primary school teacher candidates' communication skills in science and technology teaching lessons. According to the results, the majority of the teacher candidates expressed positive opinions.

Samarapungavan et al. (2023) have highlighted the use of technology-mediated, model-based inquiry lessons to engage kindergarten students in developing their models of matter. They found that students from both intervention groups showed significant improvements.

Literature Review: The study categorized under this theme involves a literature review that synthesizes existing research on the teaching and learning of properties of matter. Hadenfeldt, Liu, and Neumann (2014) have presented a systematic review of research on how students understand the matter, focusing on identifying common patterns across studies. Their findings indicated a shift in focus from categorizing students' conceptions to analyzing their progression in understanding matter over time. The review identified typical pathways through which students develop their understanding, aligning with previous frameworks on chemical reactions, physical states, atoms, molecules, and conservation. They presented a model describing students' progression in understanding matter, contributing to developing a K-12 learning progression.

Teacher education and pedagogical content knowledge: Five of the articles examined were about teacher education and their pedagogical content knowledge about teaching and learning the properties of matter. After analyzing these articles, several key findings emerged:

Pre-service teachers still hold misconceptions about density even after being taught (Harrel & Subramaniam, 2014), and elementary science teachers have difficulty understanding the significance of small particle properties of matter (Hanuscin et al., 2018). Cisterna, Bookbinder, Mikeska, & Lakhani (2022)

examined how elementary pre-service teachers viewed assessment tasks about teaching the properties of matter. The results showed that while the teachers recognized the importance of this topic for elementary classrooms, they had little experience with the subject matter. Additionally, although they understood the value of the tasks for teaching science, they struggled to link them to specific teaching practices. Mikeska, Phelps, & Croft (2017) discussed the creation and testing of assessment items to measure elementary science teachers' content knowledge. The items were developed in three science units: one of them was the properties of matter. They tried to measure elementary teachers' content knowledge. Their evidence suggests that the assessment items effectively capture variations in teachers' performance and teachers' knowledge. Avcı (2022) investigated the impact of using educational games in the Jigsaw technique on the communication skills of primary school teacher candidates in relation to the topic of properties of matter. The research found that integrating the Jigsaw technique with educational games can enhance communication skills and foster interest in science classes.

Result and Discussion

The analysis of the articles on teaching and learning about the properties of matter reveals a wide range of activities, methods, and strategies available to teachers. The analysis of the distribution of studies across different years revealed a significant concentration of research activities in 2022. This suggests a notable focus on scholarly endeavours in physics education during the last period about teaching and learning properties of matter. The majority of the studies employed qualitative methods, followed by quantitative and mixed methods. The studies targeted various study groups, including teachers, teacher candidates, primary school students, secondary school students, pre-school students, and high school students. When the study groups of the research are examined, it has been determined that some studies specifically address the needs and experiences of teachers and teacher candidates, while others focus on specific age groups within the education system. Notably, one study conducted a comprehensive literature review, contributing a valuable resource for researchers and practitioners in the field. The wide variety of study groups and research methods shows the need for further research in the field of education. These findings provide valuable insights for future research endeavours and underscore the importance of considering various study groups and employing various research methods to understand teaching and learning the properties of matter comprehensively.

These articles demonstrate a variety of activities, methods, and strategies for teaching and learning about the properties of matter. The findings suggest that approaches like the model-based science writing heuristic (M-SWH) approach (Kara & Kingir, 2022), hands-on experiments (Kácovský, 2019; Wilcox et al.,

2022), cooperative learning (Zorlu & Sezek, 2020) modeling-based teaching (Zorlu & Sezek, 2020), inquiry experiences (Samarapungavan et al., 2023; Wilcox et al., 2022), and technology-mediated lessons (Samarapungavan et al., 2023) can effectively enhance students' understanding, academic achievement, science process skills, and communication skills in the domain of properties of matter.

In summary, these findings provide evidence that implementing various teaching approaches and strategies can lead to significant gains in students' understanding, academic achievement, science process skills, and communication skills related to the properties of matter.

It is important to consider the progression of concepts, provide hands-on experiences, encourage collaborative learning, utilize modelling techniques, and leverage technology to support the teaching and learning of the properties of matter in physics. Nevertheless, despite the existence of studies focusing on specific groups in the existing literature, it is worth highlighting the absence of any research that approaches the subject matter from a holistic perspective, considering its significance and conceptual development. It is important to approach teaching about the properties of matter in a holistic way and employ a variety of methods to enhance learning outcomes. Future research in this field should continue to explore new approaches and strategies to improve education in this area.

Acknowledgement

This research is supported by the Department of Mathematical Sciences, Computer Science and Physics of the University of Udine.

References

Adams, K., & Feagin, S. (2017). Describing matter. *Science and Children, 54*(8), 52.

Akkus, A., & Doymus, K. (2022). Effect of subject jigsaw and reading writing presentation techniques on academic achievement of 7th grade science students'academic success in structure and properties of matter. *Acta Didactica Napocensia, 15*(1), 92-104.

Allen, J., & Rogers, M. P. (2015). Putting ideas on paper. *Science and Children, 53*(3), 32.

Andersson B., (1990), Pupils' conceptions of matter and its transformations (age 12–16), *Studies in Science Education, 18*, 53–85. https://doi.org/10.1080/03057269008559981

Avcı, F. (2022). Effects on primary school teacher candidates of developing and implementing jigsaw technique activities enriched with educational games in science and technology teaching lessons. *Acta Educationis Generalis, 12*(1), 41-

73.

Can, M. (2017). *Farklı kavramsal değişim stratejileriyle zenginleştirilmiş bağlam temelli yaklaşımın madde ve özellikleri konusunun anlaşılmasına etkisi* [Master›s thesis, Giresun Üniversitesi, Fen Bilimleri Enstitüsü].

Cisterna, D., Bookbinder, A. K., Mikeska, J. N., & Lakhani, H. R. (2022). Elementary preservice teachers' perceptions of assessment tasks that measure content knowledge for teaching about matter. *Journal of Science Teacher Education, 33*(8), 910-937.

Hadenfeldt, J. C., Liu, X., & Neumann, K. (2014). Framing students' progression in understanding matter: a review of previous research. *Studies in Science Education, 50*(2), 181-208.

Hanuscin, D. L., Cisterna, D., & Lipsitz, K. (2018). Elementary teachers' pedagogical content knowledge for teaching structure and properties of matter. *Journal of Science Teacher Education, 29*(8), 665-692.

Harrell, P. E., & Subramaniam, K. (2014). Teachers need to be smarter than a 5th grader: What pre-service teachers know about density. *The Electronic Journal for Research in Science & Mathematics Education, 18*(6).

Harrison, A., & Treagust, D. (2002). The particulate nature of matter: Challenges in understanding the submicroscopic world. In J. Gilbert, O. de Jong, D. F. van Driel, & J. H. van Driel (Eds.), *Chemical education: Towards research-based practice* (pp. 189–212). Dordrecht: Kluwer.

Horizon (2030). Looking ahead to challenges and opportunities. https://www.eib.org/attachments/strategies/horizon_2030_en.pdf

Kácovský, P. (2019). The visit of a thermal imaging world in one physics lesson. *Physics Education, 54*(4), 045011.

Kara, S., & Kingir, S. (2022). Implementation of the Model-Based Science Writing Heuristic Approach in elementary school science. *International Journal of Science and Mathematics Education, 20*(4), 683-703.

Krnel, D., Watson, R., & Glažar, S. A. (1998). Survey of research related to the development of the concept of 'matter'. *International Journal of Science Education, 20*(3), 257-289

Krnel, D., Watson, R., & Glažar, S. A. (2005). The development of the concept of 'matter': A cross-age study of how children describe materials. *International Journal of Science Education, 27*(3), 367-383.

Liu, X., & Lesniak, K. (2005). Students' progression of understanding the matter concept from elementary to high school. *Science Education, 89*(3), 433–450.

Liu, X. (2013). Difficulties of items related to energy and matter: Implications for learning progression in high school chemistry. *Educación química, 24*(4), 416-422.

MEB (2018a). *Fen Bilimleri dersi öğretim programı.* Milli Eğitim Bakanlığı

MEB (2018b). *Ortaöğretim Fizik dersi öğretim programı.* Milli Eğitim Bakanlığı

MIUR (2012). *Indicazioni nazionali per il curricolo della scuola dell'infanzia e del primo ciclo d'istruzione.* https://www.miur.gov.it/documents/20182/51310/

DM+254_2012.pdf/1f967360-0ca6-48fb-95e9-c15d49f18831?version=1.0&t=1480418494262 .

Mikeska, J. N., Phelps, G., & Croft, A. J. (2017). Practice-based measures of elementary science teachers' content knowledge for teaching: Initial item development and validity evidence. *ETS Research Report Series, 2017*(1), 1-72.

NGSS (n.d.). *Next Generation Science Standards*. https://www.nextgenscience.org/search-standards?keys=properties+of+matter

Oxman, A. D. (1994). Systematic reviews: checklists for review articles. *British Medical Journal, 309*(6955), 648-651. https://doi.org/10.1136/bmj.309.6955.648

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Blackwell Publishing.

Piaget J. and Inhelder B., (1974). *The child's construction of quantities: Conservation and atomism*. Psychology Press.

Pimthong, P., Yutakom, N., Roadrangka, V., Sanguanruang, S., Cowie, B., & Cooper, B. (2012). Teaching and learning about matter in grade 6 classrooms: A conceptual change approach. *International Journal of Science and Mathematics Education, 10*, 121-137.

Samarapungavan, A., Bryan, L. A., Staudt, C., Sapkota, B., Will Pinto, H. E., Broadhead, J. M., & Kimball, N. (2023). Using technology-mediated inquiry to help young learners reimagine the visible world through simple particle models. *Journal of Research in Science Teaching, 60*(2), 390-422.

Stamenkovski, S., & Zajkov, O. (2017). Seventh grade students' qualitative understanding of the concept of mass influenced by real experiments and virtual experiments. *European Journal of Physics Education, 5*(2), 6-16.

Zorlu, Y., & Sezek, F. (2020). An Investigation of the effect of students' academic achievement and science process skills application together with cooperative learning model and the modeling based teaching method in teaching science courses. *International Journal of Progressive Education, 16*(4), 135-157.

Wilcox, J., Reiter, R., Rose, A., Alberts, A., & Murano, K. (2022). Patterns That Matter: Helping Second Graders Explore Patterns of Properties of Matter and the Nature of Science. *Science and Children, 59*(3), 30-36.

The Effect of Simulation-Integrated Online POE Activities on Preservice Science Teachers' Conceptual Understanding of Simple Electric Circuits, Attitudes and Physics Self-Efficacy Beliefs

Seda Altunsoy Okvuran, Nejla Yürük

Gazi University, Turkey

This study aims to examine the effects of online POE-based teaching activities integrated with simulations on preservice science teachers' (PSTs') conceptual understanding, attitudes towards simple electrical circuits, and self-efficacy beliefs in physics. The participants of the study who were studying in a state university in Turkey constitute 29 fourth-grade PSTs. Quantitative research was conducted by the one-group pre-test-post-test experimental design. "Simple Electrical Circuits Diagnostic Test (SECDT)", "Attitude Scale Towards Simple Electrical Circuits (ASSEC)", and "Physics Self-Efficacy Scale (PSS)" were applied to PSTs as pretest and posttest. In the online courses conducted with the Zoom application, PSTs performed POE tasks, group, and class discussions online and made their observations of simple electrical circuits using PhET simulations. The results indicate that the simulation-integrated online POE activities improved the conceptual understanding of PSTs about simple electrical circuits and positively affected their attitudes towards this subject and their physics self-efficacy. The frequency of the misconceptions revealed in the pre-SECDT and post-SECDT was also provided.

Keywords: Simulation-integrated online POE, simple electrical circuits, conceptual understanding

Introduction

At present, the use of online technologies has become more widespread as individuals' access to information becomes easier and technological developments direct the education system. Especially during the pandemic period, web-based learning has become a necessity (Hadjerrouit, 2010). Various science teaching studies reveal that web-based inquiry-oriented science teaching methods and some multimedia applications can positively affect learners' conceptual understanding (Yen et al., 2011). In this regard, Predict–Observe–Explain (POE) method proposed by White and Gunstone (1992) has emerged as one of the active learning methods compatible with conceptual change pedagogy

that enables learners to construct scientific conceptual understanding (Kearney, 2004). Also, the fulfillment of the tasks of POE stages and small group and class discussions can be adapted to the online environment.

At the same time, students can become more motivated in the subject by participating in various multimedia applications such as simulation, which can be applied to real-world problems and provide students with more efficient learning opportunities (Cetinkaya & Kirilmazkaya, 2022). Interactive computer simulations especially allow students to manipulate initial conditions directly and see the effect immediately (Zacharia, 2005). Several studies on science education also indicate that web-based inquiry-oriented science teaching methods and some multimedia applications can have positive effects on learners' attitudes towards science topics and science self-efficacy (Tsai et al., 2011). Self-efficacy is a concept within social cognitive theory that proposes that an individual's belief in their own abilities is necessary for them to successfully complete tasks and achieve goals (Bandura, 2006). In this study, physics self-efficacy belief represents how capable an individual perceives themselves to be in accomplishing a task related to physics lessons. Newhouse (1990) defines attitude as positive or negative feelings towards a person, object, or subject. Accordingly, the attitude towards simple electrical circuits in this study refers to individuals' positive or negative stance towards the subject. In addition, it is known that students' high self-efficacy beliefs positively affect their learning motivation attitudes toward the subject and, consequently, academic achievements (Lin, Liang, & Tsai, 2015).

Based on these studies, we thought that POE, which is normally carried out in the classroom, can be brought into an online environment by integrating simulations into the "Observe" stage, which may positively affect conceptual understanding, attitudes, and physics self-efficacy of preservice science teachers (PSTs). In other science teaching studies, the stages of POE were commonly executed by integrating them into simulations. Students predicted, observed, and explained the outcome of the simulation results (James et al., 2022; Rutten et al., 2015). However, this study integrated simulations into the 'Observe' stage of POE in line with the online learning environment. Thus, during the Predict stage, the PSTs will have the opportunity to encounter different ideas in group discussions and to feel the limitations of their own ideas by questioning the reasons behind them before moving on to observation. However, in the various studies, there is a greater emphasis on investigating the effects of learners' self-efficacy beliefs related to internet usage and participation in web-based learning (Salame & Samson, 2019; Tsai et al., 2023). Therefore, there needs to be more research regarding the effects of web-based inquiry learning on attitudes and self-efficacy toward science topics. For this reason, in this study, we aimed to examine the effect of the simulation integrated-online POE method on PSTs' conceptual

understanding of simple electric circuits subject, in which students often have misconceptions, have learning difficulties, and include abstract concepts, and on their attitudes towards this subject and physics self-efficacy.

Alternative Conceptions in Simple Electric Circuits

Conceptual learning studies in science education have indicated that students at different educational levels and countries hold a wide range of similar misconceptions about simple electrical circuits (Peşman & Eryılmaz, 2010; Tahir et al., 2020). These misconceptions arise from the fact that students often use the terms energy, current, voltage, and potential difference interchangeably (Engelhardt & Beichner, 2004). Table 1 explains the common misconceptions identified in the literature by Peşman (2005) and Peşman and Eryılmaz (2010), with the abbreviations used in this study are provided in parentheses:

Table 1. Explanations of Common Misconceptions in Electrical Circuits

Misconceptions	Explanation
Short -circuit Misconception (SM)	involves disregarding wires without electrical devices in the analysis of an electrical circuit.
Clashing Current Model (CCM)	in which positive and negative electricity from the power supply meet at an electrical device, and their clash causes it to power up.
Shared Current Model (SCM)	posits that electrical devices are believed to distribute the electrical current equally among them.
Attenuation Model (ATM)	suggests that when flowing through a circuit in a single direction, an electrical current is progressively diminished due to its consumption by devices within the circuit.
Sequential Reasoning (SR)	it is believed that any alteration at a certain point in an electrical circuit impacts the circuit in the direction of the current flow rather than in the reverse direction.
Local Reasoning (LR)	is the assumption that when the current reaches each junction point in the circuit, regardless of what happens, it is divided into as many equal parts as the number of junctions.
Power Supply as a Constant Current Source (PCCS)	it is assumed that power supplies deliver a constant electrical current to a circuit instead of providing electrical energy.
Sink Model (SNM)	it is believed that a single wire connection between a power supply and an electrical device is sufficient to power the device.
Empirical Rule Model (ERM)	it is believed that a bulb's brightness decreases as its distance from the battery increases.
Parallel Circuit Misconception (PCM)	involves the incorrect belief that adding more resistors in parallel raises the total resistance.
Current Flow as Water Flow (CFWF)	in thinking of the movement of charges in a wire as the movement of water in pipes.

Research Questions

In this study in which simulation-integrated online POE activities were applied to PSTs, the following research questions are sought to be answered:

1. Is there a statistically significant difference between PSTs' conceptual understanding mean scores on the Simple Electrical Circuits Diagnostic Test (SECDT) before and after the teaching activities?
2. Is there a statistically significant difference between the mean scores of PSTs on the Attitude Scale Towards Simple Electrical Circuits (ASSEC) before and after the teaching activities?
3. Is there a statistically significant difference between the mean scores of PSTs on the Physics Self-Efficacy Scale (PSS)" before and after the teaching activities?

Method

This quantitative research was conducted using a one-group pretest-posttest experimental design. The participants of this study, who were enrolled at a state university in Turkey, comprised 29 PSTs in their fourth year. Participants were selected through convenience sampling based on the willingness of PSTs to participate in teaching activities. The teaching activities were implemented in the course "Misconceptions in Science Teaching" and lasted for five weeks, two hours per week. As data collection tools, "Simple Electrical Circuits Diagnostic Test (SECDT)", "Attitude Scale Towards Simple Electrical Circuits (ASSEC)", and "Physics Self-Efficacy Scale (PSS)" were applied to PSTs before and after the teaching activities. The characteristics of the data collection tools, and the purposes of their implementation are presented in Table 2.

Table 2. Characteristics and Purposes of Data Collection Tools

Test/Scale	The purpose of implementation	Characteristics of the tool
Simple Electrical Circuits Diagnostic Test (SECDT)	to determine the alternative concepts of PSTs about simple electrical circuits and the change in their conceptual understanding	developed by Peşman (2005) and adapted for university students (Arı, Peşman, & Baykara, 2016). a 3-tier test 12 items, Cronbach alpha 0,69 (calculated using correct answers). reveals 11 misconceptions
Attitude Scale Towards Simple Electrical Circuits (ASSEC)	to assess PSTs' attitudes toward the topic of simple electrical circuits	developed by Taşlıdere and Eryılmaz (2012) for high school students. 24 items, 5 sub-dimensions Cronbach alpha 0,93 5-point Likert scale
Physics Self-Efficacy Scale (PSS)	to assess PSTs' physics self-efficacy beliefs	developed by Çalışkan (2007) for university students. 24 items, 4 sub-dimensions Cronbach alpha 0,94 5-point Likert scale

The first stage of the SECDT test items consists of typical multiple-choice questions. The second tier asks for the reason behind the answer in the first tier, and the third tier questions the participants' confidence in the answers given in the first two tiers. In this study, PSTs' conceptual understanding scores were calculated based on their correct answers to all three tiers of SECDT. Therefore, even if the PSTs provided scientifically accurate responses to the first two stages, an answer of "not sure" for the third tier was not considered scientifically correct. If a PST marked options indicating a misconception but was unsure at the third tier, their answer was not considered a misconception. Additionally, we calculated the frequency values of misconceptions obtained from the pretest and posttest of SECDT.

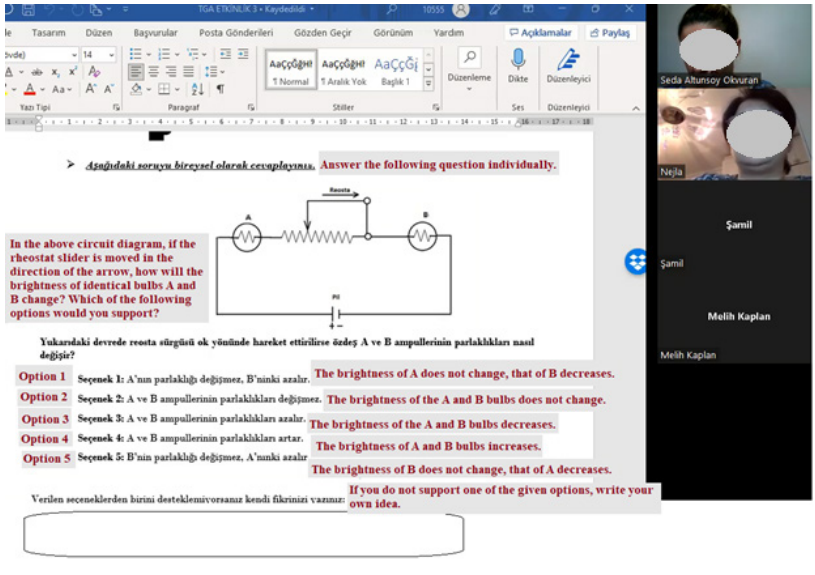
Simulation-integrated Online POE-based Activities

We preferred the Zoom meeting application, one of the popular applications used for online learning, to carry out activities based on the simulation integrated POE method online. We have prepared five activity forms based on POE to eliminate alternative concepts that are frequently encountered in the literature. At the Zoom meeting, the researcher first shared the activity form with the PSTs and projected it on the screen. The simulation-integrated online POE activities were conducted in the following stages:

Predict

The Predict stage starts with a simple electrical circuit diagram question provided at the beginning of the activity form that reveals the alternative conceptions held by PSTs. They first make individual predictions about the situation given in the diagram. The researcher asked PSTs to choose one of the theories given in the activity form when making their predictions. Then, PSTs with different predictions were assigned to discussion groups by creating breakout rooms. During group discussions, each PST tried to convince others of the correctness of their own prediction by explaining why they disagreed with the opinions of their groupmates and by explaining their reasons in detail.

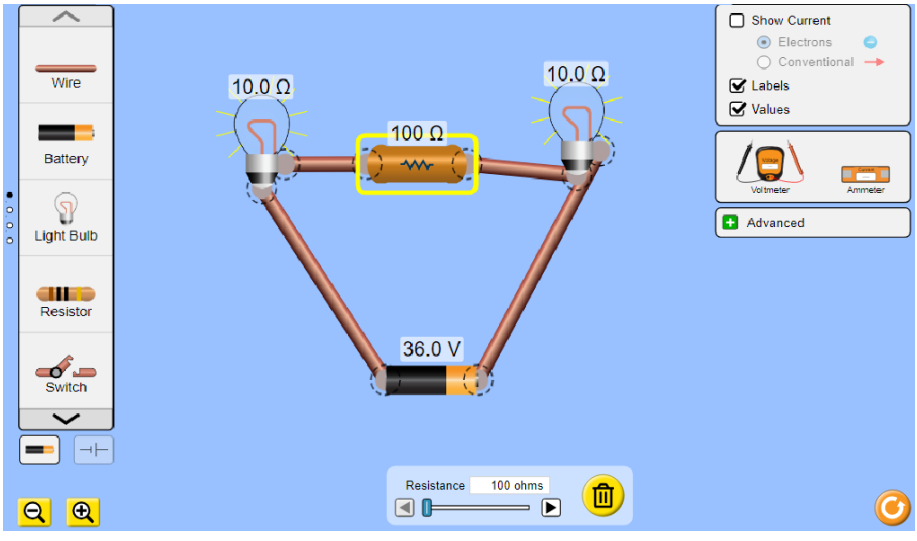
Figure 1. The Discussion Group in a Breakout Room



Observe

At the Observe stage, all the rooms were closed, and the PSTs returned to the main meeting room. Then, the researcher performed the same simple electrical circuit diagram as given at the beginning of the activity form with the circuit construction kit of the PhET interactive simulation site. Afterward, PSTs simulated the same circuit on their personal computers using PhET. The researcher asked them to express how their predictions and observations differed.

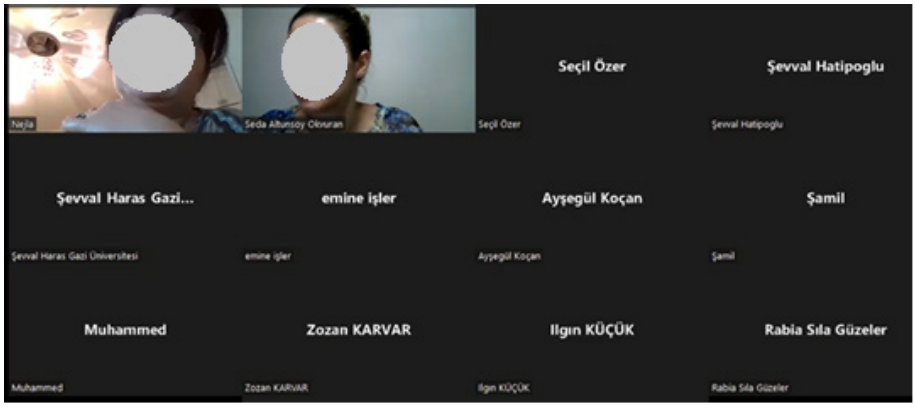
Figure 2. Setup of the Simple Electrical Circuit Diagram Using PhET



Explain

The PSTs discussed the reasons for the observation results as a class and were expected to compare their previous predictions with the data they obtained from the observation. Afterward, the researcher presented scientific explanations about why the previous predictions of the PSTs were not correct based on the observation results.

Figure 3. Class Discussion in the Explain Stage



All online meetings were recorded using the video recording option in Zoom so that PSTs had the chance to repeat the lesson by accessing this recording whenever they needed to.

Role of Researcher in Teaching Activities

During group discussions, the researcher took turns participating in breakout rooms and asked some questions to help PSTs realize their preconceptions and the reasons behind these concepts. At the same time, with these questions, researcher aimed to make PSTs dissatisfied with their scientifically inaccurate concepts. For example, if a PST using a statement such as, "*the charges that constitute the current leave the positive pole of the battery, pass through A and B bulbs, and return to the negative pole of the battery"* in a circuit where A and B bulbs are connected in series, the researcher posed a counter question as "*However, in our daily lives, when we press a switch, we observe that both bulbs light up simultaneously. According to your assumption, shouldn't we have seen the A bulb light up first, followed by the B bulb?"* In the Explain stage, the researcher guided the PSTs and encouraged them to consider all possible reasons for the observation and to provide alternative interpretations. In this way, the researcher allowed the PSTs to construct their own concepts before presenting scientific explanations. In addition, the researcher clarified concepts that PSTs did not understand well in the Explain stage and reconstructed the simple electrical circuit shown in the diagram using PhET if requested by the PSTs.

Results

In the study, the paired samples t-test was conducted to analyze the effects of the teaching method on PSTs' mean scores of conceptual understandings, attitudes towards simple electrical circuits, and physics self-efficacy beliefs to answer the research questions. Table 3 presents the t-test results of comparing PSTs' SECDT, ASSEC, and PSS pretest and posttest mean scores.

Table 3. T-Test Results for PSTs' Pretest and Posttest Mean Scores of SECDT, ASSEC, and PSS

Dependent variable	Time	N	M	SD	df	t	p	Cohen's d
SECDT	pretest	29	4,24	2,81	28	9,53	,000*	,76
	posttest	29	9,07	1,53				
ASSEC	pretest	29	88,48	14,70	28	3,71	,004*	,33
	posttest	29	94,07	14,61				
PSS	pretest	29	89,97	14,32	28	2,39	,023*	,17
	posttest	29	94,65	12,81				

*There is a significant difference at the p <0,05 level.

According to the results of the analysis in Table 3, there was a statistically significant increase in SECDT mean scores from pretest (M=4.24, SD=2.81) to posttest (M=9.07, SD=1.53), t (28)=9.53, p< .05; ASSEC mean scores from

pretest (*M*=88.48, *SD*=14.70) to posttest (*M*=94.07, *SD*=14.61), *t* (28)= 3.71, *p*< .05; and PSS mean scores from pretest (*M*=89.97, *SD*=14.32) to posttest (*M*=94.65, *SD*=12.81), *t* (28)= 2.39, *p*< .05. Cohen's d effect size values in parentheses for SECDT (.76), ASSEC (.33), and PSS (.17) indicated a large effect size for all variables (Cohen, 1988).

We also concluded that the simulation-integrated online POE method is more effective in eliminating some misconceptions about simple electrical circuits. The frequency values (f) of the misconceptions obtained from the pretest and posttest of SECDT are given in Figure 4.

Figure 4. Frequency of Misconceptions Revealed in the Pretest and Posttest of SECDT.



Figure 4 illustrates that PSTs were unable to eliminate the "Local Reasoning (LR)" and "Shared Current Model (SCM)" misconceptions at the end of the teaching activities. However, fewer PSTs had the other six misconceptions after the online teaching activities than before. The most common misconception that was eliminated is the "Short Circuit Misconception (SM)."

Discussions and Conclusions

The research findings indicate that simulation-integrated POE activities significantly enhanced PSTs' conceptual understanding, attitudes towards simple electrical circuits and their self-efficacy in physics. Accordingly, in this study, similar results were obtained with studies found that web-based inquiry teaching practices in science teaching and specifically the PhET simulated POE method, had a positive effect on learners' conceptual understanding of science subjects and motivational beliefs such as attitudes and self-efficacy (Cetinkaya & Kirilmazkaya, 2022; Tsai et al., 2011; Yen et al., 2011). Through the online teaching activities in this study, conceptual explanations, visuals, and experiments became clear and repeated, and each PST was provided to become aware of their preconceptions by participating in the group discussions. Consequently, they had the opportunity to reflect on their previous conceptions and reconcile them with their current ones, thereby enhancing their ability to quickly apply and integrate what they learned into new situations (Yen et al., 2011). As a result, they were able to construct their concepts in a scientifically accurate manner.

At the same time, this finding can be interpreted to imply that the integration of PhET simulations into POE contributed to making the lessons more interesting rather than monotonous, therefore enhancing PSTs' learning and questioning skills and improving their conceptual understanding (Alamsyah et al., 2021). This improvement, in turn, could positively impact their attitudes toward the subject matter and their self-efficacy beliefs (Lin et al., 2015). We also think that PhET simulations enable the complex circuits to be set up in a more practical and simple way, contributing positively to the efficient use of time and especially the self-efficacy of PSTs who may have difficulty in constructing experimental setups using real experiment materials.

The misconception of "Local Reasoning (LR)" arises from a scientific inaccuracy about how current is generated in a circuit. Conceptual change of this misconception is challenging as it requires a shift in students' ontological presuppositions about the current. When the frequency change of misconceptions from the pretest to the posttest of SECDT is analyzed, in the posttest, the frequency of PSTs who answered "I am sure of my answers to the first two tiers" to the third tier of questions measuring the local reasoning misconception increased compared to the pretest. Therefore, some PSTs were identified as having local reasoning misconceptions in the posttest. In addition, there were also PSTs who could not eliminate the misconception of "Local Reasoning (LR)" and "Shared Current Model (SCM)" from pretest to posttest. However, most PSTs in both situations experienced increased self-efficacy after participating in the teaching activities. The increased self-efficacy contributed to their confidence in their preconceptions and resistance to changing their misconceptions. In the literature, this situation emphasizes that self-efficacy beliefs can affect conceptual change in two ways. High self-efficacy enables learners to change their preconceptions and trust their abilities to reach a better conceptual understanding; on the other hand, it can also cause learners to resist changing their own concepts (Sinatra & Mason, 2008).

Acknowledgement

We thank all preservice science teachers who voluntarily participated in this study.

References

Alamsyah, M., Marhento, G., Siburian, M. F., Astuti, I. A. D., & Bhakti, Y. B. (2021). Application of blended learning with Edmodo based on POE learning model to increase students understanding of science concepts. In *Journal of Physics: Conference Series*: Vol. 1806, (pp 012121). IOP Publishing.

Arı, Ü., Peşman, H., & Baykara, O. (2017). Sorgulamaya dayalı öğretimde rehberlik düzeyinin bilimsel süreç becerileriyle etkileşimi. *Bartın Üniversitesi Eğitim*

Fakültesi Dergisi, 6 (1), 304-321. doi:10.14686/buefad.263895

Bandura, A. (2006). Guide to the construction of self-efficacy scales. In F. Pajares & T. Urdan (Eds.). *Self-efficacy beliefs of adolescents,* (Vol. 5., pp. 307-337). Greenwich, CT: IAP - Information Age Publishing.

Cetinkaya, U., & Kirilmazkaya, G. (2022). The effect of POE method with PhET simulation on primary school student's science attitudes and success: Greenhouse effect. *Education Quarterly Reviews, 5*(4), 350-360. doi:10.31014/aior.1993.05.04.597

Cohen, J. W. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, NJ: Lawrence Erlbaum Associates.

Çalışkan, S. (2007). *Problem çözme stratejileri öğretiminin fizik başarısı, tutumu, özyeterliği üzerindeki etkileri ve strateji kullanımı.* (Unpublished doctoral dissertation). Dokuz Eylül Üniversitesi, İzmir, Turkey.

Engelhardt, P. V., & Beichner, R. J. (2004). Students' understanding of direct current resistive electrical circuits. *American Journal of Physics, 72*(1), 98-115. doi:10.1119/1.1614813

Hadjerrouit, S. (2010). Developing web-based learning resources in school education: A user-centered approach. interdisciplinary. *Journal of E-Learning and Learning Objects, 6*(1), 115-135. Retrieved from https://www.learntechlib.org/p/44777/

James, N. M., Kreager, B. Z., & LaDue, N. D. (2022). Predict-observe-explain activities preserve introductory geology students' self-efficacy. *Journal of Geoscience Education*, *70*(2), 238-249. doi:10.1080/10899995.2021.1906593

Kearney, M. (2004). Classroom use of multimedia-supported predict-observe-explain tasks in a social constructivist learning environment. *Research in Science Education, 34*(4), 427-453. doi:10.1007/s11165-004-8795-y

Lin, T. C., Liang, J. C., & Tsai, C. C. (2015). Conceptions of memorizing and understanding in learning, and self-efficacy held by university biology majors. *International Journal of Science Education, 37*(3), 446-468. doi:10.1080/09500693.2014.992057

Newhouse, N. (1990) Implications of attitude and behavior research for environmental conservation. *The Journal of Environmental Education, 22*(1), 26-32. doi:10.1080/00958964.1990.9943043

Peşman, H. (2005). *Development of a three-tıer test to assess ninth grade students' misconceptions about simple electric circuits*. (Unpublished master's thesis). Middle East Technical University, Ankara, Turkey.

Peşman, H., & Eryılmaz, A. (2010). Development of a three-tier test to assess misconceptions about simple electric circuits. *The Journal of Educational Research*, *103*(3), 208-222. doi:10.1080/00220670903383002

Rutten, N., van der Veen, J. T., & van Joolingen, W. R. (2015). Inquiry-based whole-class teaching with computer simulations in physics. *International Journal of Science Education*, *37*(8), 1225-1245. doi:10.1080/09500693.2015.1029033

Salame, I. I., & Samson, D. (2019). Examining the implementation of PhET simulations

into general chemistry laboratory. *International Journal of Environmental and Science Education, 14*(4), 207-217. Retrieved from http://www.ijese.net/makale/2116.html

Sinatra, G. M., & Mason, L. (2008). Beyond knowledge: Learner characteristics influencing conceptual change. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 560-582). New York: Routledge.

Tahir, F. M., Nasri, N. M., & Halim, L. (2020). The effectiveness of predict-observe-explain-animation (POE-A) strategy to overcome students' misconceptions about electric circuits concepts. *Learning Science and Mathematics*, *15*, 1-15. Retrieved from https://myjms.mohe.gov.my/index.php/lsm/article/view/9457/5748

Taşlıdere, E., & Eryılmaz, A. (2012). Basit elektrik devreleri konusuna yönelik tutum ölçeği geliştirilmesi ve öğrencilerin tutumlarının değerlendirilmesi. *Journal of Turkish Science Education*, *9*(1), 31-46. Retrieved from https://www.tused.org/index.php/tused/article/view/419/357.

Tsai, C. C., Chuang, S. C., Liang, J. C., & Tsai, M. J. (2011). Self-efficacy in Internet-based learning environments: A literature review. *Journal of Educational Technology & Society*, *14*(4), 222-240. doi:10.2307/jeductechsoci.14.4.222

White, R., & Gunstone, R. (1992). *Probing understanding*. London and New York: The Falmer Press.

Yen, H. C., Tuan, H. L., & Liao, C. H. (2011). Investigating the influence of motivation on students' conceptual learning outcomes in web-based vs. classroom-based science teaching contexts. *Research in Science Education*, *41*(2), 211-224. doi:10.1007/s11165-009-9161-x

Zacharias C. Z. (2005) The impact of interactive computer simulations on the nature and quality of postgraduate science teachers' explanations in physics. *International Journal of Science Education, 27*(14), 1741-1767. doi:10.1080/09500690500239664.

Students' Reasoning on the Dependence of Energy and Structure in Chemical Reactions

Benjamin Pölloth

Eberhard Karls Universität Tübingen, Department for Chemistry Education, Germany

The dependence of energy on chemical structures is central to the field of chemistry. However, empirical studies show that students' conceptual understanding of energy is rather low. The aim of this study is to investigate which mental resources high school students use to reason about kinetic and thermodynamic aspects of chemical reactions. In a qualitative interview study, 38 high school students from 11th to 13th grade in 16 focus groups were asked to explain the role of energy in the reaction of hydrogen and chlorine to a tutoring student. Qualitative content analysis of the interviews showed that students used a variety of cognitive resources to rationalise energy changes. However, many students focused on technical terms and argued at a system level rather than at the level of the chemical structure of particles. Only very rarely did students link energetic aspects to changes in chemical bonds. The driving force of the reaction seemed to be beyond the scope of many students and was most dominantly explained by the octet rule. The results suggest the need to develop new learning and teaching approaches to strengthen the mental link between energetic and structural aspects of chemical reactions.

Keywords: Conceptual understanding, energy, semi-structured interviews

Introduction

One of the central questions in chemistry is why certain compounds react the way they do. A meta-perspective of recent chemical research shows that systematically tracing reactions back to molecular structures and their energies is the key to understanding reactivity (Goodwin, 2007; Pölloth et al., 2022). While this approach has been tedious for many years, the establishment of quantum chemical methods in recent years has revolutionized the way reactivity can be analyzed: Computational chemistry has become a standard method to directly access energies of almost any structure (Seeman & Tantillo, 2022). Thus, Jensen (2007) concludes: "Chemistry is knowing the energy as a function of the nuclear coordinates."

In contrast, empirical research shows that learners' understanding of energy is low and barely connected to other central chemical concepts (Podschuweit & Bernholt, 2018). Energy is rarely understood as a result of particle interactions

(Macrie-Shuck & Talanquer, 2020). Therefore, competence in the concept area of energy increases marginally during school career in Germany (Bernholt et al., 2020), and students rarely discuss energetic aspects for the rationalizing chemical processes (Abell & Bretz, 2018; Eckhard et al., 2022).

This gap between the scientific understanding of chemistry and students' understanding of the role of energy indicates the need to develop new teaching and learning approaches. Therefore, it is necessary to identify the cognitive resources that students draw on to explain energetic aspects in chemistry. These existing cognitive resources may form the basis for connecting new ideas (diSessa, 2018). Although it is unlikely that learners have a single coherent concept of a complex scientific idea such as energy, students do have different mental resources related to energetic aspects in chemistry (Hammer et al., 2005). In doing so, which of these mental resources students activate depends critically on the context of the problem (diSessa, 2018).

Research Questions

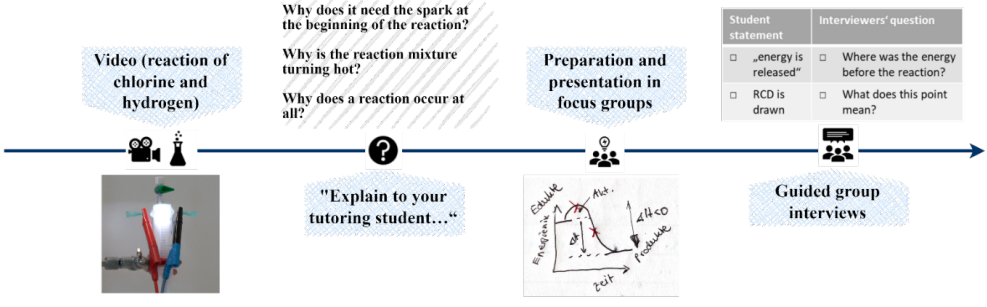
Hence, in this study high school students were asked to reason on energetic aspects of a specific chemical reaction. Their activated mental resources were analysed in detail to answer these research questions:

1. Which cognitive resources do students draw on to explain kinetic and thermodynamic aspects of a chemical reaction?
2. In how far do students argue with potential energy as a function of the chemical structure?

Methods

In the qualitative study design, problem-oriented guided focus groups interviews were conducted with 38 German high school students from 11th to 13th grade (Pölloth et al., 2023). In the first part of the study, students watched a short video clip of the reaction of a mixture of chlorine and hydrogen gas, that was ignited by an electric spark. In focus groups, the high school students worked on an explanation for a tutoring student (see Figure 1). They presented their results to the interviewer who took the role of the tutoring student. Follow-up questions about the groups' reasoning were then discussed using an interview guide. This approach makes it possible to distinguish clearly between ideas activated by the problem itself and those introduced by the interviewer.

Figure 1. Sequence of the interview study with a screenshot of the experimental video, the guiding questions for the interview, a student drawn reaction coordinate diagram from the focus group time and a short excerpt of the interview guide (from left to right).



Qualitative data analysis (Kuckartz, 2016) was conducted on the transcribed interviews. Main categories were formed deductively based on the research questions, and sub-categories were derived inductively from the data material. Parts of the data material were re-coded by a second person with a good intercoder reliability (K = 0.81). In addition, students' self-concept of ability, scientific interest and school grades in chemistry were assessed in a short paper-pencil survey. The results indicate a broad distribution of self-concept of ability and interest across the sample.

Results and Discussion

The structured analysis of the 16 focus groups allowed the identification of archetypical argumentation patterns of students. Two representative case studies are used here to illustrate the findings. Group Pink was a group of two students with a medium scientific interest and self-concept of ability in chemistry. Group Gold consisted of two highly interested students with a high self-concept of ability in chemistry and a third student, with intermediate scores on both dimensions.

Function of the spark – activation energy

Group Pink explained the need of the "spark" as follows:

> "[The spark] is sort of the trigger, the activation energy, and so it acts as the reaction start, so that it reacts with each other at all."

Students did not use further cognitive resources in their initial explanation. Figure 2 illustrates this argumentation by the solid connecting lines. The explanation focused mainly on the macroscopic scale, while the term "activation energy" was used more as a symbolic black box term. Many students equated

the activation energy and the “spark” - the energy added - on the same level. This suggests that students do not readily associate the term activation energy with the submicroscopic scale but rather with the macroscopic level. In contrast, the Gold group gave the following initial explanation:

> „That means [the reaction mixture] needs a certain energy level, perhaps that the molecules move faster, that they can collide faster.”

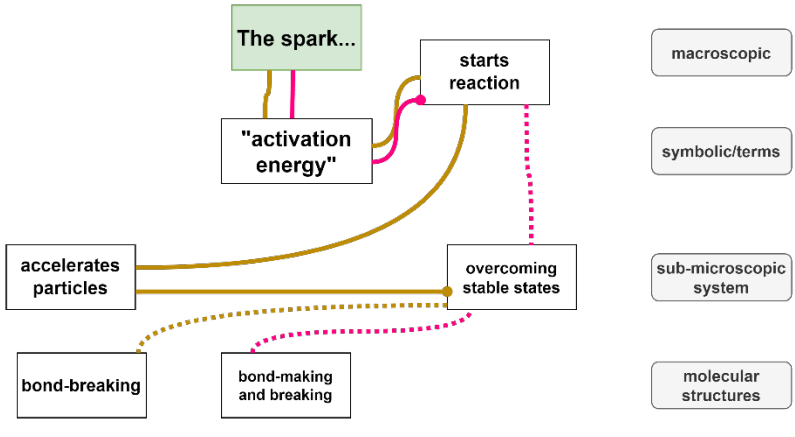
While this argument is at the sub-microscopic level, it understands particles as a kind of entity in a system. However, it does not argue at the productive level of chemical thinking, which is chemical structure. Thus, in the second interview phase, the students were explicitly asked about the relationship between the energy required and chemical bonds. The resources activated by these questions are connected to the dotted bonds in Figure 2, as in the example of Group Gold:

> „2: So it means [the reaction] needs energy solely for breaking the old bonds and the new bonds just forms like that on its own.“

Thus, the main reason why students did not argue with chemical structure does not seem to be a lack of knowledge. Rather, from a resource-oriented perspective, it seems that students are not used to activate resources about bond-making and -breaking when arguing about energy. Similarly, Group Pink was able to refer to chemical bonds, although the argumentation was not very concise:

> „Maybe it [the spark] is doing something to the atom or molecule or something that they’re bonding or reacting with each other that they wouldn’t normally do. [...] Maybe they get mixed up somehow....”

Figure 2. Visualization of the argumentation patterns of Group Pink and Gold on the function of the spark.



Solid lines indicate that students used these arguments in their initial unguided explanation. Arguments connected with dashed lines were only activated during the guided interview. Overviews for all focus

groups can be found in Pölloth et al. (2023).

Similar results were found for all the groups (for details see Pölloth et al., 2023). Argumentation on a macroscopic level and the extensive use of symbolic terms were dominant. Only a few groups argued on a sub-microscopic level, but most of them did not refer to chemical structures or bonds at all. About half of the groups mentioned bond cleavage as the main factor for activation energy when asked directly.

Origin of the heat - thermodynamics

Similar argumentation patterns were found, when students were asked to rationalize the origin of the heat. The term "exothermic" played the crucial role, as the example of Group Pink shows (Figure 3):

> "This is an exothermic reaction. And thereby simply heat is generated."

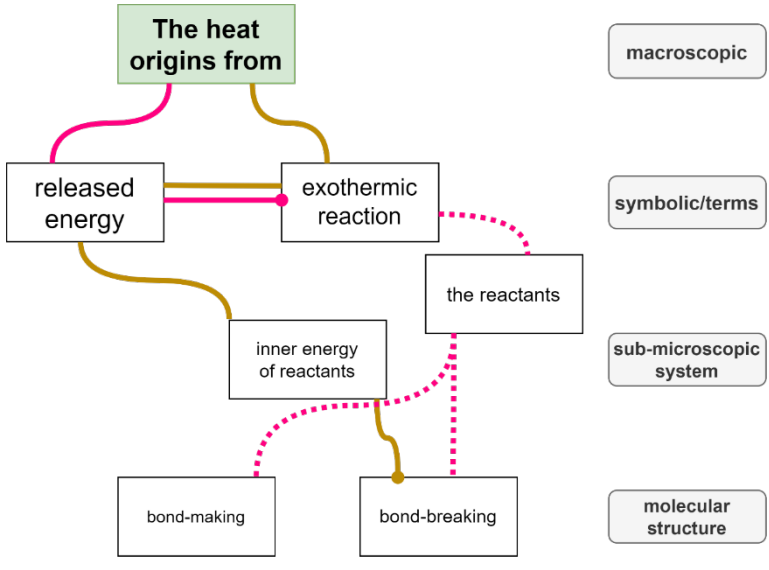
In fact, cause and effect are often confused in this context. Exothermic is not used to describe the heat balance of a reaction, but to explain it.

Group Gold gave an explanation that included the level of chemical structures:

> "2: Yes, [heat is released] for example, when bonds break or so, when a molecule reacts, for example [...] because the bond is held together by energy."

Interestingly, the same student who correctly argued above about the need for energy to break bonds here shows an understanding of exothermic bond breaking (Galley, 2004; Hunter et al., 2022). These two examples suggest that the students' understanding of bond breaking does not seem to be a stable and coherent "misconception". It seems more appropriate to argue from a resource-oriented perspective, as the student seem to activate resources in different ways depending on the context.

Figure 3. Visualization of the argumentation patterns of Group Pink and Gold on the origin of the heat.



Solid lines indicate that students used these arguments in their initial unguided explanation. Arguments connected with dashed lines were only activated during the guided interview. Overviews for all focus groups can be found in Pölloth et al. (2023).

In the whole sample, the resources of bond-making were activated by only one group to rationalize the origin of heat. Even when asked directly, only a quarter of the students were able to relate bond-making and energetic aspects.

Why does a reaction occur – driving force

The last guiding question put the focus on the driving force of the reaction. Very dominantly groups activated resources related to the octet rule (Figure 4), as Group Pink does:

> "Yeah, they just want the octet rule, um, that just the last shell is full."

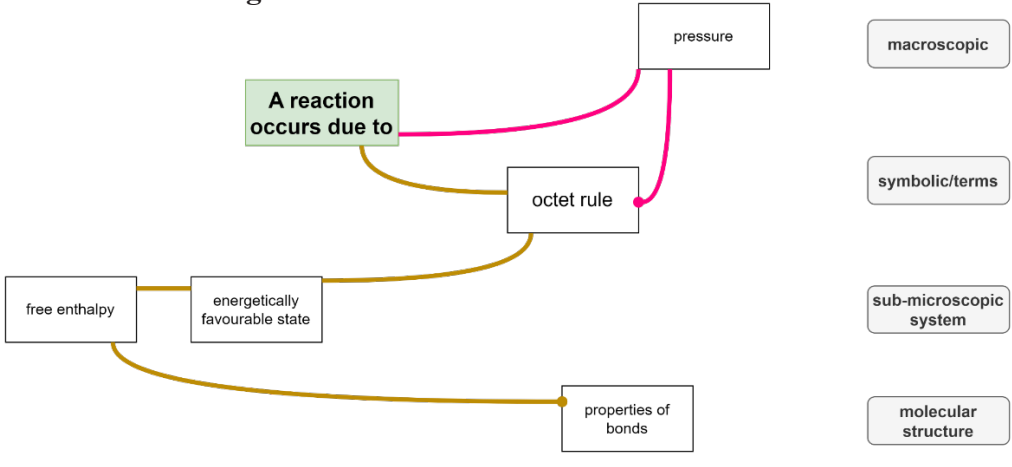
This reasoning is not meaningful in the given context because all reagents and products in this reaction (H_2, Cl_2, HCl) fulfill the octet rule. Nevertheless, students continue to use the octet rule even after this problem has been discussed. Activation of the octet rule was found to be very dominant.

Group Gold used a wide range of knowledge in their argumentation (Figure 4):

> "Reactions always take place [...] voluntarily, if at the end of the reaction there is an energetically favorable state for the substances. And that is the case here, because obviously HCl is energetically more favorable than if one, um, hangs around alone."

They argued for the formation of energetically favorable states and even referred to the structural level. However, there seems to be a missing piece in the students' argumentation (e.g. differences in the stability of bonds). Hence, the dominant activation of the octet rule could also origin from a lack of alternatives. It seems that arguments on the driving force of reactions are out of scope for most of the groups in the sample.

Figure 4. Visualization of the argumentation patterns of Group Pink and Gold on the driving force of the reaction.



Solid lines indicate that students used these arguments in their initial unguided explanation. Arguments connected with dashed lines were only activated during the guided interview. Overviews for all focus groups can be found in Pölloth et al. (2023).

Limitations

As in every qualitative study, results are not representative and can be specific for the interviewed students. As group interviews were performed, students influence each other. Hence, it is not possible to isolate a single students' argumentation. The study does not allow to distinguish if the ideas students articulate are mainly adopted from by school lessons or other influences, or if they were constructed by the students themselves.

Discussion and Implications

The analysis of the students' statements revealed a wide variety of mental resources for chemical energetics. Technical terms such as "exotherm" or "activation energy" were often taken as a comprehensive explanation of activation energy and thermodynamics. While these terms allowed students to quickly find answers to the first two guiding questions, they all had great difficulty arguing about the driving force. Therefore, most students activated resources related to the octet rule, even though they were not productive for the problem.

Many students did not argue at the particle level in the unguided phase of the interview. Even when asked directly about the submicroscopic level, most students argued in terms of the whole system rather than discussing structural changes in individual particles. In fact, it does not seem clear what we are talking about when we talk about the submicroscopic scale. It should be clearly emphasized that the structural level is the most productive level when discussing chemistry (National Research Council, 2012, p. 121).

Particularly, the central concept of breaking and making chemical bonds was activated seldomly to argue about energetic changes (Nahum et al., 2007). For many students, chemical bonds and structures appear to be weakly connected to the energetic aspects of chemical reactions. As outlined above, the activation of mental resources about structure is a key factor for productive scientific reasoning about energy in chemistry.

Therefore, it is desirable for high school energy education to emphasize that knowing the energy of a given structure is the key to understanding chemical reactions (Jensen, 2007). In line with the claims of other scientists, this could include the following relationships:

- All atoms interact through attractive and repulsive electrostatic forces (National Research Council, 2012)
- In a chemical bond and in stable molecules, attractive and repulsive forces are in equilibrium (Nahum et al., 2007)
- Chemical reactions are rearrangement of atoms; the minimization of potential energy through the formation of more stable bonds is one driving force of reactions (Cooper & Klymkowsky, 2013)

Acknowledgement

The author thanks the “Fonds der chemischen Industrie” (FCI) for financial support; Dominik Diekemper and Stefan Schwarzer for collaboration in the project; and all students participating in the study.

References

Abell, T. N., & Bretz, S. L. (2018). Dissolving Salts in Water: Students’ Particulate Explanations of Temperature Changes. *Journal of Chemical Education*, *95*(4), 504–511. https://doi.org/10.1021/acs.jchemed.7b00845

Bernholt, S., Höft, L., & Parchmann, I. (2020). Die Entwicklung fachlicher Basiskonzepte im Chemieunterricht – Findet ein kumulativer Aufbau im Kompetenzbereich Fachwissen statt? [Central scientific concepts in chemistry education—Does a cumulative development of conceptual competences take place?]. *Unterrichtswissenschaft*, *48*(1), 35–59. https://doi.org/10.1007/s42010-019-00065-4

Cooper, M. M., & Klymkowsky, M. W. (2013). The trouble with chemical energy: Why understanding bond energies requires an interdisciplinary systems approach. *CBE Life Sciences Education*, *12*(2), 306–312. https://doi.org/10.1187/cbe.12-10-0170

diSessa, A. A. (2018). A Friendly Introduction to "Knowledge in Pieces": Modeling Types of Knowledge and Their Roles in Learning. In G. Kaiser, H. Forgasz, M. Graven, A. Kuzniak, E. Simmt, & B. Xu (Eds.), *ICME-13 Monographs. Invited Lectures from the 13th International Congress on Mathematical Education* (pp. 65–84). Springer International Publishing. https://doi.org/10.1007/978-3-319-72170-5_5

Eckhard, J., Rodemer, M., Bernholt, S., & Graulich, N. (2022). What Do University Students Truly Learn When Watching Tutorial Videos in Organic Chemistry? An Exploratory Study Focusing on Mechanistic Reasoning. *Journal of Chemical Education*, *99*(6), 2231–2244. https://doi.org/10.1021/acs.jchemed.2c00076

Galley, W. C. (2004). Exothermic Bond Breaking: A Persistent Misconception. *Journal of Chemical Education*, *81*(4), 523. https://doi.org/10.1021/ed081p523

Goodwin, W. (2007). Scientific Understanding after the Ingold Revolution in Organic Chemistry. *Philosophy of Science*, *74*(3), 386–408. https://doi.org/10.1086/522358

Hammer, D., Elby, A., Scherr, R. E., & Redish, E. F. (2005). Resources, framing, and transfer. In J. Mestre (Ed.), *Transfer of Learning from a Modern Multidisciplinary Perspective* (pp. 89–120). Information Age Publishing.

Hunter, K. H., Rodriguez, J.-M. G., & Becker, N. M. (2022). A Review of Research on the Teaching and Learning of Chemical Bonding. *Journal of Chemical Education*, *99*(7), 2451–2464. https://doi.org/10.1021/acs.jchemed.2c00034

Jensen, F. (2007). *Introduction to computational chemistry* (2nd edition). John Wiley & Sons.

Kuckartz, U. (2016). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung* (3., überarbeitete Aufl.). *Grundlagentexte Methoden*. Beltz. http://nbn-resolving.org/urn:nbn:de:bsz:31-epflicht-1117809

Macrie-Shuck, M., & Talanquer, V. (2020). Exploring Students' Explanations of Energy Transfer and Transformation. *Journal of Chemical Education*, *97*(12), 4225–4234. https://doi.org/10.1021/acs.jchemed.0c00984

Nahum, T. L., Mamlok-Naaman, R., Hofstein, A., & Krajcik, J. (2007). Developing a new teaching approach for the chemical bonding concept aligned with current scientific and pedagogical knowledge. *Science Education*, *91*(4), 579–603. https://doi.org/10.1002/sce.20201

National Research Council. (2012). *A Framework for K-12 Science Education*. National Academies Press. https://doi.org/10.17226/13165

Podschuweit, S., & Bernholt, S. (2018). Composition-Effects of Context-based Learning Opportunities on Students' Understanding of Energy. *Research in Science Education*, *48*(4), 717–752. https://doi.org/10.1007/s11165-016-9585-z

Pölloth, B., Diekemper, D., & Schwarzer, S. (2023). What resources do high school students activate to link energetic and structural changes in chemical reactions? – A qualitative study. *Chemistry Education Research and Practice*, *24*(4), 1153–1173. https://doi.org/10.1039/D3RP00068K

Pölloth, B., Häfner, M., & Schwarzer, S. (2022). At the same time or one after the other? – Exploring reaction paths of nucleophilic substitution reactions using historic insights and experiments. *CHEMKON*, *29*(2), 77–83. https://doi.org/10.1002/ckon.202100060

Seeman, J. I., & Tantillo, D. J. (2022). Understanding chemistry: From "heuristic (soft) explanations and reasoning by analogy" to "quantum chemistry". *Chemical Science*, *13*(39), 11461–11486. https://doi.org/10.1039/d2sc02535c

Investigating the International Applicability of the MAVerBE Analysis Model

Dennis Dietz[1], Ian Belinchón[1], Claus Bolte[1], Mariona Espinet[2]

[1]Freie Universität Berlin, Berlin, Germany

[2]Universitat Autònoma de Barcelona, Barcelona, Spain

Learning is successful when learners develop knowledge structures that are linked in as many ways as possible; knowledge structures that are not passively anchored but can be actively used in real-life situations. To assess the success of learning processes, the Model for the Analysis of the Linkage of Term Elements (German acronym: MAVerBE) can be used to examine the linking performance of students regarding three dimensions: 1. vertical linkage level, 2. horizontal (cross-curricular) linkage and 3. scientific correctness. MAVerBE has been used several times in recent years to investigate the linking performance of students in Germany regarding the interdisciplinary energy concept. The aim of this study was to test MAVerBE in the Catalan context and to gain a first insight into the linking performance of Catalan students regarding the energy concept. For this purpose, we asked 109 students from the Barcelona metropolitan region to write an essay on the energy concept. We analysed these essays using MAVerBE in a qualitative content analysis procedure. The results show that the students in our sample have acquired a significant, predominantly scientifically correct knowledge on the energy concept. However, they seldomly formulate complex statements that correspond to a higher level of vertical linkage.

Keywords: Knowledge-linking, conceptual understanding, energy concept.

Introduction

From the perspective of constructivist learning theory, successful learning is characterized by the fact that new elements of knowledge are linked to existing elements of knowledge in as many ways as possible (e.g., Renkl, 1996). For this reason, linking is a quality criterion for knowledge (e.g., de Jong & Ferguson-Hessler, 1996). The correlation between successful knowledge acquisition and linking performance has not only been described theoretically, but has also been intensively studied and empirically proven (e.g., Neumann et al., 2008; Knobloch et al., 2013; Podschuweit et al., 2016; Podschuweit & Bernholt, 2020). One option for assessing linked – and therefore successful – learning in student statements is to use the Model for the Analysis of the Linkage of Term Elements (acronym: MAVerBE, in German: Modell zur Analyse der Vernetzung von Begriffselementen; Dietz & Bolte, 2022; Dietz, 2023). In recent years, we

have used MAVerBE intensively to investigate the linking performances of students in Germany in the context of the energy concept (Dietz & Bolte, 2021; 2022; Dietz, 2023). In this paper, we present a study in which we examined the performance of students from the Barcelona metropolitan region in relation to the linking of term elements of the energy concept. The aims of this study were a) to test MAVerBE in another linguistic context and b) to gain a first empirical insight into the linking performance of Catalan students within the energy concept.

Theoretical Framework

For reasons of cognitive economy, humans form terms in order to be able to deal with the complexity of their environment (Aebli; 1981; 2003; Schnotz, 1994). During term formation, term elements are linked together in a network (Dörner, 1976; Aebli, 1981; 2003). Within this network, the term elements are linked to each other in the form of complexion hierarchies ("has-as-part-relations") and abstraction hierarchies ("is-a-part-of-relations") (Dörner, 1976). When an individual reconstructs a term – for example, when asked to explain the term – the hierarchical structure of the term networks (consisting of linked term elements) becomes temporarily visible (Aebli, 1981). For this reason, it is possible to analyse the latent construct "knowledge-linking" by examining the ability of individuals to bring term elements together in a meaningful way (Dietz & Bolte, 2021; 2022; Dietz 2023).

Over the last two decades, various models have been developed in Germany to systematically investigate students' ability to link term elements. These models include the Model of vertical Linkage (MvL) by Fischer, Glemnitz, Kauertz and Sumfleth (2007) and the Model of Hierarchical Complexity (MHC-C) adapted for Chemistry lessons by Bernholt and Parchmann (2011). Although both the MvL and the MHCC show great theoretical similarities (Woitkowski et al., 2011), the practical research application of both models leads to significantly different assessments of students' linking performance (Nehring et al., 2017). In addition, both the MvL (Fischer et al., 2007) and the MHCC (Bernholt & Parchmann, 2011) focus exclusively on the description of vertical linking performance and do not consider the dimensions of horizontal (cross-curricular) linkage and scientific correctness.

For this reason, Dietz and Bolte (2021; 2022) developed the Model for the Analysis of the Linkage of Term Elements (German acronym: MAVerBE). With MAVerBE it is possible to examine the linking performance of students regarding the following three dimensions: 1. vertical linkage level, 2. horizontal linkage and 3. scientific correctness (Fig. 1).

The first dimension of the MAVerBE "vertical linkage level" describes the complexity with which term elements are linked (Dietz & Bolte, 2021; 2022;

Dietz, 2023). With a higher vertical linkage level, an increasing number of term elements are connected with each other in an increasingly complex way. At the first vertical linkage level a maximum of two term elements are brought into a meaningful context using auxiliary verbs such as "to be" or "to have". At the second vertical linkage level (level 2: ungrounded relations), statements are considered in which at least three term elements are linked in the form of causalities, dependencies, conditions, or process descriptions. At the higher vertical linkage levels, the statements of the second vertical linkage level are combined with each other, e.g., in the form of causal chains (level 3: interconnected relations) or justifications (level 4: grounded relations). At the highest vertical linkage level, which we refer to as "multi-perspective generalisations", MAVerBE consider those students' explanations in which they link sub-concepts of the energy concept (energy forms, energy transformation, energy transfer, energy devaluation, energy conservation, and entropy (Liu & McKeough, 2005; Neumann et al., 2013; Poggi et al., 2017)) and explain the linkage using at least one example. The special linking performance is that the students use term elements at a very high level of abstraction (Dörner, 1976) and link the term element hierarchies temporarily formed during the reconstruction of the term with each other (Aebli, 1981).

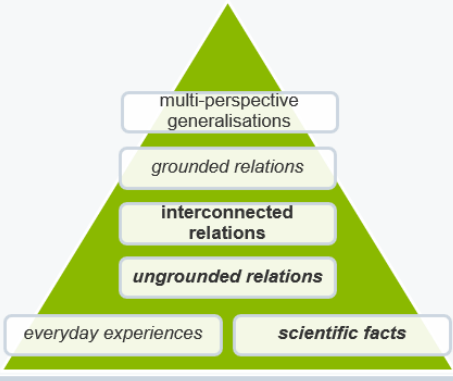
The second dimension of MAVerBE focusses on horizontal (cross-curricular) linkage (Fig. 1). To investigate horizontal linkage, the term elements in the statements from the analysis of the vertical linkage level are assigned to a subject (biology, chemistry, or physics) or a combination of subjects (biology/chemistry, etc.) based on their presence in the respective curricula of our federal state Berlin (Dietz & Bolte, 2021; 2022; Dietz, 2023). Term elements of the energy concept that are not listed in any of the curricula are categorised as "scientific-technical" if they fulfil the criteria of the taxonomy of scientific words (Wellington & Ireson, 2008).

The third dimension of MAVerBE takes into account the scientific correctness of the statements from the analysis of vertical linkage level (Fig. 1). In the "implicitly correct" category, we consider students' statements that are inaccurate with respect to content or (scientific) language demands (Dietz & Bolte, 2021; 2022; Dietz, 2023).

Fig. 1. MAVerBE contains a three-dimensional category system for analysing students' linking performance

MAVerBE – Model for the Analysis of the Linkage of Term Elements

1st dimension: vertical linkage level



The vertical linkage level defines units of analysis.

Note: The dimension "vertical linkage level" combines the **Model of vertical Linkage** (Fischer et al., 2007, elements shown in bold) and the *Model of Hierarchical Complexity adapted for Chemistry lessons* (Bernholt & Parchmann, 2011, elements shown in italics). Additionally, we optimized the category definitions of both models and developed an own definition for the highest category.

2nd dimension: horizontal linkage

biology (B)	chemistry (C)	physics (P)
biology/chemistry (BC)	biology/physics (BP)	chemistry/physics (CP)
biology/chemistry/ physics (BCP)	energy (E)	scientific-technical (NT)

Identification of term elements within a unit of analysis and assignment to a category on the basis of curricula

3rd dimension: scientific correctness

explicitly correct	implicitly correct	incorrect

Analysis of scientific correctness within a unit of analysis

Over the past years, we successfully piloted and (afterwards) used MAVerBE to investigate the linking performance of students within the energy concept (Dietz & Bolte, 2021; 2022; Dietz, 2023). We chose the energy concept as the object of investigation for two reasons: 1) The energy concept is extremely complex, which can be recognized, among other things, by the fact that it can be described by the following six sub-concepts: energy form, energy transformation, energy transfer, energy devaluation, energy conservation, and entropy (e.g., Liu & McKeough, 2005; Neumann et al., 2013; Poggi et al., 2017; Dietz, 2023). Thus, links between term elements of the energy concept can be expected at different levels of complexity. 2) The energy concept is relevant in all natural science subjects (e.g., Eisenkraft et al., 2014), so that horizontal (cross-curricular) linkages of term elements are also well expectable.

The studies on the linking performance of students in Germany have shown that although they build up significant knowledge structures on the energy concept, they predominantly link term elements of the energy concept at a low vertical linkage level (Dietz & Bolte, 2021; 2022; Dietz, 2023). The linking performance

of students regarding the energy concept can be increased if science is taught in an integrated rather than subject-differentiated way (Dietz, 2023). Regardless of the form of science teaching, in the context of the analyses of the scientific correctness of the students' statements it was notable that these were partly due to language barriers. For instance, students frequently misinterpreted the sub-concept of energy conservation as energy production, as the word "erhalten" can mean both "to conserve" and "to receive" in German (Dietz, 2023).

Based on our research on students' understanding of the energy concept in Germany, we were interested in a) how MAVerBE works in other linguistic contexts and b) how well students in other countries are able to link term elements of the energy concept. Due to a cooperation between the "Didactics of Chemistry" working group at the Freie Universität Berlin and the "Language and Science Education-LIEC" working group at the Universitat Autònoma de Barcelona, we were able to investigate the following research questions as part of the study presented here:

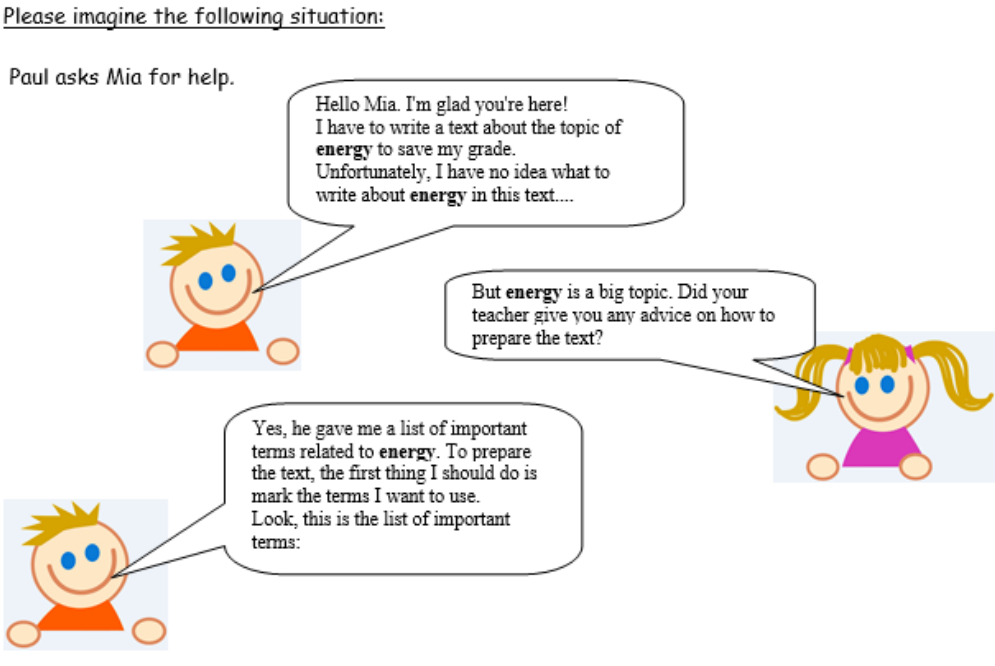
1. How well does the analysis of linking performances regarding the energy concept work with MAVerBE in the Catalan context?
2. To what extent are students from the Barcelona metropolitan region able to link term elements of the energy concept in a complex and scientifically correct way?

Method

In order to answer these research questions, we asked students in the fourth grade of secondary school (equivalent of 10th grade in Germany) in the Barcelona metropolitan region to write an essay on the energy concept during a regular school lesson (45 min). As essay writing tasks pose high linguistic demands on the students, the students were allowed to decide for themselves whether to write the essays in Catalan or Spanish – depending on in which language they felt they could express themselves better. To avoid overwhelming the students with the task, we provided them a list of 26 term elements of the energy concept, whose relevance was validated in preliminary studies through curriculum analyses and teacher surveys in Berlin (N = 106) (Dietz & Bolte, 2021; 2022; Dietz, 2023). The task was also embedded in an everyday context to motivate the students to write an essay on the energy concept (Fig. 2).

For an initial exploration of the extent to which MAVerBE can be used in the Catalan context and to examine to what extent the students are able to link term elements of the energy concept with each other, we have focused in this study on the dimensions of "vertical linkage level" and "scientific correctness". All analyses were conducted with software support using MAXQDA software (VERBI software, 2019).

Fig. 2. The survey instrument (translated to Catalan language by I. Belinchón)



endothermic	*kinetic energy*	*metabolism*	*energy forms*
system	*photosynthesis*	*reaction energy*	*work*
energy flow	*chemical energy*	*electric energy*	*nutrition*
chemical reaction	*energy transformation*	*particles*	*potential energy*
cell respiration	*activation energy*	*heat*	*light*
energy conservation	*mitochondrion*	*exothermic*	*motion*
	energy content	*temperature*	

Now it's your turn. Put yourself in Paul's shoes and write a text about energy.

1. First, mark the terms in the box (above) that you want to use for your text on **energy**.
2. If you think of other terms that you would like to use in your text about **energy**, write them in the empty box.

3. Write a detailed text on the topic of **energy** using the marked terms.

Results

In October 2021, we asked in six classes at two schools in the Barcelona metropolitan region a total of 109 students to write an essay about the energy

concept. The students (male: 62, female: 42, non-binary: 5) were on average 15.2 years old at the time of the survey. Three of the 109 essays had to be excluded from further analysis because the students either did not write anything at all or the essays were not related to the energy concept. The remaining 106 essays are on average 97.1 ± 93 words long. 86 essays were written in Catalan and 20 in Spanish.

We formed on average 11.8 ± 7.9 units of analysis per essays from student statements. These units of analysis are distributed across all categories of the "vertical linkage level" dimension. Examples for each category of the vertical linkage level dimension are shown in Table 1.

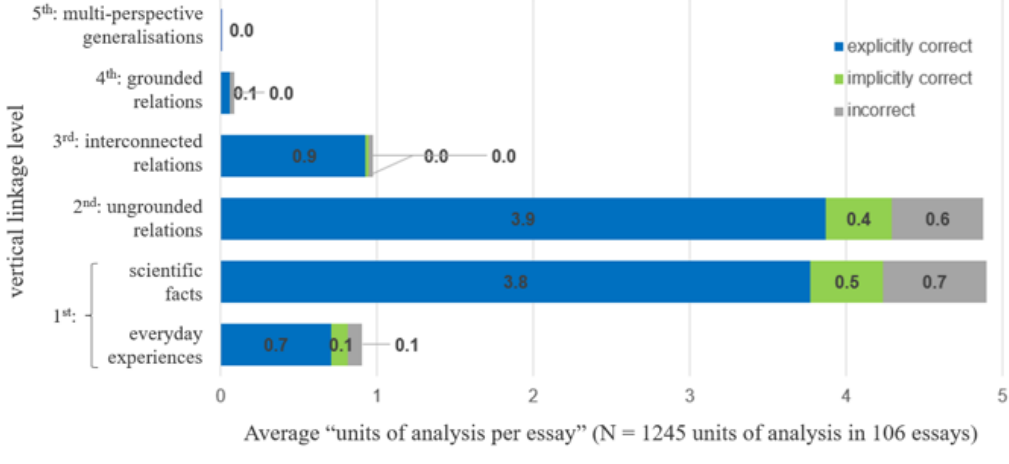
Table 1. Selected coding examples for the dimension vertical linkage level

Ia. Everyday Knowledge	
"Una energia és algo que fa que tinguis força tu i més."	Energy is something that gives you strength and more.
Ib. Scientific Fact	
"L'energia és la capacitat de realitzar un treball."	Energy is the ability to do work.
II. Ungrounded Relations	
"Tot moviment necesita una energia per realitzarla"	Every movement requires energy to happen.
III. Interconnected Relations (underlined)	
"Les plantes es nodreixen mitjançant la fotosíntesi de la mateixa manera que nosaltres realitzem el metabolisme per aconseguir energia."	Plant nutrition happens through photosynthesis in the same way that metabolism works as a means to gain energy in us [the human body].
IV. Grounded Relations	
"quan llenças alguna cosa al terra amb forca anirà molt més ràpid que si tu la deixas caura,	If you throw something strongly against the floor, it will go much faster than if you just let it drop,
per que tu li has donat energía perque fosi més rapid"	because you have provided it with energy to make it go faster
V. Multi-perspective Generalisations	
"No es pot crear ni destruir, sinó que només cambia de forma.	It [Energy] cannot be created or destroyed; it can only change form.
Per exemple si llencem un vago desde sobre d'una muntanya russa	For example, if we release a wagon from the top of a roller-coaster
l'energia del vago és totalment potencial,	the energy of the wagon is totally potential,

	pero quan comença a baixar disminueix l'energia potencial,	when it starts coming down, the potential energy diminishes,
	pero no significa que desapareixi,	but it does not mean it disappears,
	si no que es transforma en energía cinetica	but that it transforms into kinetic energy
	que provoca que més tard amb l'acomulació d'aquesta puigui puyar una pujada més petita que l'anterior.	which causes that, later, with the accumulation of the kinetic energy, it [the wagon] can climb a smaller rise than the previous one.
	Poc a poc amb la fricció de les rodes amb les vies la fricció provoca que una part d'energía cinética i potencial es transforma en calorífica."	Slowly, the friction of the wheels with the tracks causes that a part of the kinetic and potential energy transforms into thermal energy.

The average category occupancy in the "vertical linkage level" dimension – differentiated according to their assessment in the "scientific correctness" dimension – is shown in Figure 3.

Figure 3. Average number of units of analysis per essay for the different categories describing the dimension "vertical linkage level" – differentiated according to the dimension "scientific correctness"



Most units of analysis have to be assigned to the first (scientific facts) or the second vertical linkage level (ungrounded relations) (Fig. 3). The students' essays contain hardly any statements that correspond to the highest – the fourth or fifth – vertical linkage levels. With regard to the scientific correctness dimension, it should be noted that the majority of the statements can be assessed as explicitly correct, both in overall assessment (88.1 % explicitly or implicitly correct assessed statements) and with regard to the individual vertical linkage levels.

Discussion

The application of MAVerBE to essays in Catalan or Spanish was possible without any problems. As in the German language, the coding of the vertical linkage level can be orientated towards certain words or sentence structures. For instance, statements using "through" or "if... then..." (causalities), "the more... the more/less" (dependencies) and "without..." (conditions) can be clearly assigned to the category "ungrounded relations" at the second vertical linkage level, if they are not linked to other statements of these category.

In principle, the results – particularly in view of the length of the essays and the high percentage of scientifically correct statements – allow the interpretation that the students in the sample have acquired significant knowledge about the energy concept. However, it should be noted that the students in the sample were only able to link term elements of the energy concept at a low vertical linkage level. This observation is consistent with the results of studies on the vertical linking performance of students in Germany regarding the energy concept (Dietz & Bolte, 2021; 2022; Dietz, 2023).

Interestingly, our analyses revealed that there are also obstacles to a scientifically correct understanding of the energy concept in the Catalan and Spanish languages. Similar to German students (see Theoretical Framework), the students surveyed had problems correctly describing the sub-concept "energy conservation". In contrast to the German students, however, the Catalan students of our survey misinterpreted the concept of "energy conservation" not as "energy generation" but as "energy saving". These findings – both those from this study and from the studies on students' understanding of the energy concept in Germany (Dietz & Bolte, 2021; 2022; Dietz, 2023) – emphasise the obstacles to learning that everyday language has for the scientifically correct learning of scientific concepts (e.g., Rincke, 2011; Gieske et al., 2022).

Outlook

In Future Studies At Freie Universität Berlin, We Will Focus On The Important Topic Of "Climate Change". We Are Going To Use An Adapted Version Of Maverbe To Investigate The Extent To Which Students Are Able To Link Relevant Term Elements From The Topic "Climate Change" In A Complex And Scientifically Correct Way. We Are Particularly Interested In Investigating Whether And To What Extent The Form Of Science Teaching (In Germany, More Traditionally Differentiated Into The Subjects Of Biology, Chemistry, And Physics, Or In Many European And English-Speaking Countries Typical Integrated Science Lessons) Has An Influence On Students' Linking Performance Regarding The Topic "Climate Change". To Realise This Research Project,

It Would Be Exciting To Work Together With National And International Cooperation Partners. If You Are Interested, Please Contact Us At The Following E-Mail Address: Dennis.Dietz@Fu-Berlin.De

Acknowledgements

We would like to thank the contribution of Alba Montalban and Núria Janer for opening their secondary schools in the metropolitan area of Barcelona and collaborate in the process of data collection. This research was partially funded by the Spanish Ministry of Science and Innovation (PID2022-138166NB-C22b) within the framework of the ACELEC consolidated research group recognized by AGAUR (2021SGR00647).

References

Aebli, H. (1981). *Denken: das Ordnen des Tuns*. *Band 2: Denkprozesse*. Klett-Cotta.

Aebli, H. (2003). *Zwölf Grundformen des Lehrens* (12. Auflage). Klett-Cotta.

Bernholt, S., & Parchmann, I. (2011). Assessing the complexity of students´ knowledge in chemistry. *Chem. Educ. Res. Pract.*, 12, 167–173. https://doi.org/10.1039/C1RP90021H

De Jong, T. & Ferguson-Hessler, M. G. M. (1996). Types and Qualities of Knowledge. *Educational Psychologist, 31(2)*, 105–113. https://doi.org/10.1207/s15326985ep3102_2

Dietz, D., & Bolte, C. (2021). Analysis of Linking Knowledge within the Concept of Energy in Student Essays. In: G. S. Carvalho, A. S. Afonso, & Z. Anastácio (Eds.), *Fostering Scientific Citizenship in an Uncertain World* (pp. 465–473). European Science Education Research Association, virtuelle Konferenz 2021. University of Minho, Braga, Portugal.

Dietz, D., & Bolte, C. (2022). Multidimensional Analysis of Knowledge-Linking within the Concept of Energy in Student Essays, *NorDiNa*, 18(3), 353–368. https://doi.org/10.5617/nordina.9284

Dietz, D. (2023). *Vernetztes Lernen im fächerdifferenzierten und integrierten naturwissenschaftlichen Unterricht aufgezeigt am Basiskonzept Energie. Eine Studie zur Analyse der Wirksamkeit der Konzeption und Implementation eines schulinternen Curriculums für das Unterrichtsfach „Integrierte Naturwissenschaften 7/8“.* Logos.

Dörner, D. (1976). *Problemlösen als Informationsverarbeitung.* Kohlhammer.

Eisenkraft A., Nordine, J., Chen, R. F., Fortus, D., Krajcik, J., Neumann, K., & Scheff, A. (2014) Introduction: Why Focus on Energy Instruction? In: R. F. Chen, A. Eisenkraft, D. Fortus, J. Krajcik, K. Neumann, J. Nordine, & A. Scheff (alle Hrsg.), *Teaching and Learning of Energy in K – 12 Education* (S. 1–11). Springer, Cham. https://doi.org/10.1007/978-3-319-05017-1_1

Fischer, H. E., Glemnitz, I., Kauertz, A., & Sumfleth, E. (2007). Auf Wissen aufbauen –

Kumulatives Lernen in Chemie und Physik. In: E. Kircher, R. Girwidz, P. Häußler (Hrsg.), *Physikdidaktik. Theorie und Praxis* (S. 657–678). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-34091-1_22

Gieske, R., Streller, S., & Bolte, C. (2022). Transferring language instruction into science education: Evaluating a novel approach to language- and subject-integrated science teaching and learning. *RISTAL, 5,* 144–162. https://doi.org/10.2478/ristal-2022-0111

Knobloch, R., Sumfleth, E., & Walpuski, M. (2013). Förderung der Qualität fachinhaltlicher Schüleräußerungen in experimenteller Kleingruppenarbeit im Chemieunterricht. *ZfDN, 19,* 347–373.

Liu, X., & McKeough, A. (2005). Developmental growth in students' concept of energy: Analysis of selected items from the TIMSS database. *Journal of Research in Science Teaching, 42(5)*, 493–517. https://doi.org/10.1002/tea.20060

Nehring, A., Päßler, A., & Tiemann, R. (2017). The Complexity of Teacher Questions in Chemistry Classrooms: an Empirical Analysis on the Basis of Two Competence Models. *International Journal of Science and Mathematics Education, 15,* 233–250. https://doi.org/10.1007/s10763-015-9683-9

Neumann, K., Lau, A., Fischer, H. E., & Sumfleth, E. (2008). Vertikale Vernetzung und kumulatives Lernen im Chemie- und Physikunterricht. In E.-M. Lankes (Hrsg.), *Pädagogische Professionalität als Gegenstand empirischer Forschung* (S. 141–151). Waxmann.

Neumann, K., Viering, T., Boone, W. J., & Fischer, H. E. (2013). Towards a learning progression of energy. *Journal of Research in Science Teaching, 50(2),* 162–188. https://doi.org/10.1002/tea.21061

Podschuweit, S., Bernholt, S., & Brückmann, M. (2016). Classroom learning and achievement: how the complexity of classroom interaction impacts students´ learning. *Research in Science & Technological Education, 34(2),* 142–163. https://doi.org/10.1080/02635143.2015.1092955

Podschuweit, S., & Bernholt, S. (2020). Investigating Network Coherence to Assess Students' Conceptual Understanding of Energy. *Education Sciences., 10(4),* 103, 1–20. https://doi.org/10.3390/educsci10040103

Poggi, V., Miceli, C., & Testa, I. (2017). Teaching energy using an integrated science approach. *Phys. Educ., 52,* 1–9. https://doi.org/10.1088/1361-6552/52/1/015018

Renkl, A. (1996). Träges Wissen: Wenn Erlerntes nicht genutzt wird. *Psychologische Rundschau, 47,* 78–92.

Rincke, K. (2011). It's Rather like Learning a Language: Development of talk and conceptual understanding in mechanics lessons. *International Journal of Science Education, 33(2),* 229–258. https://doi.org/10.1080/09500691003615343

Schnotz, W. (1994). *Aufbau von Wissensstrukturen. Untersuchungen zur Kohärenzbildung beim Wissenserwerb mit Texten.* Beltz.

VERBI Software. (2019). MAXQDA 2020 [computer software]. Berlin, Germany:

VERBI Software. Available from maxqda.com

Wellington, J., & Ireson, G. (2008). *Science Learning, Science Teaching.* New York: Routledge.

Woitkowski, D., Riese, J., & Reinhold, P. (2011). Modellierung fachwissenschaftlicher Kompetenz angehender Physiklehrkräfte. *ZfDN, 17,* 289–313.

A Qualitative Model of the Evaporation Process for Science Education

Paulo Vitor Teodoro[1], Paulo Salles[2], Ricardo Gauche[2]

[1]Federal University of Uberlândia, Brazil.

[2]University of Brasília, Brazil

It is important understanding how students learn chemical concepts, including it has been a great concern for researchers of chemical/science education. This community is making efforts to identify the most important misunderstandings and develop strategies to overcome conceptual problems. Qualitative Reasoning (QR) has great potential for building conceptual models that can be useful for chemical education. QR is an area of Artificial Intelligence that uses the symbolic reasoning to represent mathematical functions without numbers and with explicitly modeled relationships of causality. This paper describes a qualitative model of the evaporation process, to support understanding of interactions between change of the physical state of matter and energy exchanges. The model support explanations about changes in the physical state of the matter from the kinetic energy point of view. Finally, we discuss the potential of QR articulate models for science education.

Keywords: Evaporation process. qualitative reasoning. science teaching.

Introduction

What will happen to the physical state of the water molecules if we increase the temperature? What will happen to the aggregation state of the molecules if we decrease the temperature of the water? Any Brazilian student in a secondary school should be able to answer this question, given that the physical states of water are largely used to build up concepts on the relation between physical states of matter and the exchange of energy.

Teachers try to bring this knowledge into everyday life, as it explains the functioning of a several natural phenomena concerning to science teaching. However, students can hardly give a causal account of the mechanisms working on the evaporation process behaviour. Communication between the student and the teacher about transition of physical states by means of energy exchange is difficult.

Textbooks are widely used in Brazilian schools (for example, see in Oliveira & Petrucci-Rosa, 2016), but they communicate a simplistic view of science/chemistry and do not contribute for the development of fundamental concepts,

including those about the transition of physical states of matter (for example, Souza, 2019). Thus, Souza (2019) identified the potential of qualitative modeling, based on Qualitative Reasoning (QR), that students can think about system dynamics and the development of systems thinking.

QR is an area of Artificial Intelligence (AI) that uses the symbolic reasoning to represent mathematical functions without numbers and with explicitly modeled relationships of causality. QR has great potential for developing conceptual models that show the structure and functioning of the system (for an example, in water resources management, see Salles et al. 2006; or in environmental education, see Souza et al. 2017). Mustapha et al. (2002) show that QR models in chemical education was explored in a system for simulating a chemistry laboratory. Using the ontology provided by the Qualitative Process Theory (QPT) (Forbus, 1984), models can support secondary school students in qualitative analysis of transition of physical states involved in evaporation process. Here we describe the implementation of a qualitative model for students to understand the structure and behaviour of the evaporation process.

The Modelling Process

The QPT approach (Forbus, 1984) was chosen because it allows for the explicit representation of objects, quantities, possible values, processes and qualitatively distinct states in which the system can be found. The models were implemented in DynaLearn (Bredeweg *et al.*, 2013). However, we can also implement the models in Garp (Bredeweg, 1992). Relevant properties of the objects are represented as quantities.

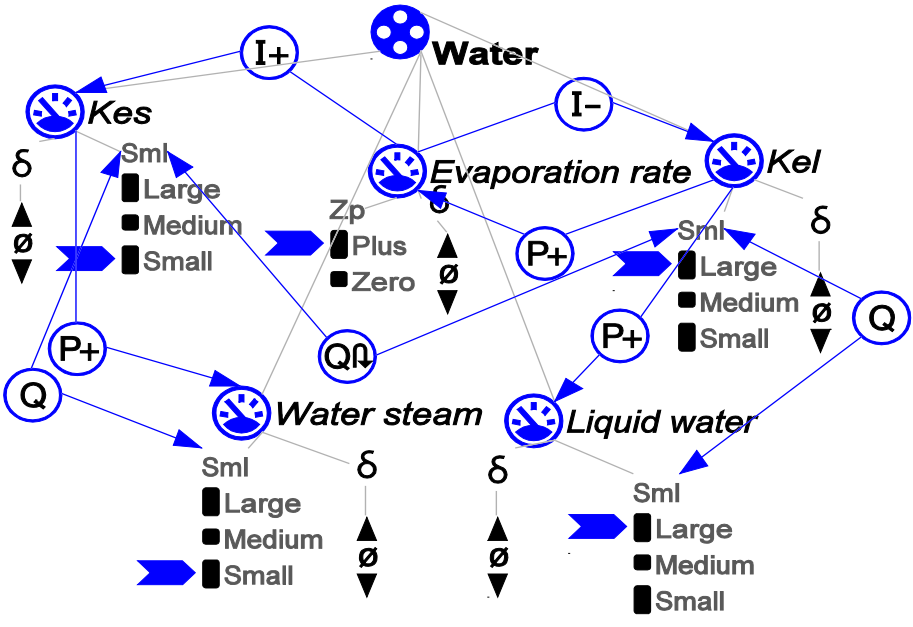
Qualitative values of quantities are given by the two variables: <magnitude, derivative>. The magnitude represents the amount of a quantity (for example, the concentration of a substance can be small), and the derivative represents its direction of change (that is, the quantity is increasing). Quantity spaces (QS) used in this model include {Small, Medium, Large} for the quantity Kinetic Energy of Steam, Kinetic Energy of the Liquid, Water Steam and Liquid Water; {zero, positive} for the quantity Evaporation Rate. For all the quantities, derivatives draw on the QS {+, 0, –}. Causality is represented by direct influences (I+ or I-) and qualitative proportionalities (P+ or P-). In QPT, the former represents processes, the primary cause of change. If, for example, I+ (A, B), then the value of B (the rate of the process) is added to the magnitude of A and, as a result, A starts to increase (if B is positive) or decrease (if B is negative).

Proportionalities propagate to other quantities changes caused by processes. In this case, the causal link is established via the derivatives. If P+ (A, B) and the derivative of B changes, it causes the derivative of A to change in the same direction. For example, if B is increasing ([derivative of B = +]) then it causes A to be increasing ([derivative of A = +]).

Simulations and Results

This model is about one of the most important phenomena of the water cycle: the process of evaporation. In the prepared Scenario (Figure 1), we can show that the Evaporation Rate influences the Kinetic Energy of Steam (KES), which in turn affects the amount of Water Vapor in the environment. Furthermore, the Evaporation Rate also affects the Kinetic Energy of the Liquid (KEL), and consequently the amount of liquid water.

Figure 1. Model scenario "Evaporation with energy changes"



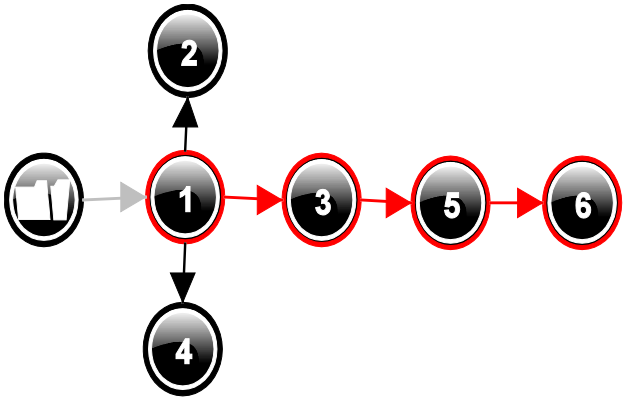
In that scenario (Figure 1), the process is represented by the *Evaporation Rate*, which influences the *KES* with I+, and I- the *KEL*. This process causes changes in the entire system, since it initiates the chain of causality. So, when the *Evaporation Rate* is positive and the direct influence is positive (I+), the amount of *KES* will increase, up to a certain limit, since the system is being controlled by the balancing feedback mechanism with a P+ (*Evaporation Rate* ← *KEL)* which indicates that the processes related to these quantities are controlled and do not continue indefinitely. And if KES increases, the amount of water vapor increases in the same direction. It is important to point out that, as KES and Steam of water are connected by the correspondence Q, these two quantities vary with the same qualitative values, the same magnitude.

Another relationship we have is about the *Evaporation Rate* and the *KEL*, connected by the direct influence (I-). In this way, when the *Evaporation Rate* is positive, the amount of *KEL* decreases up to a certain limit, because with the feedback effect, it will stabilize at a certain moment. And if the *KEL* decreases,

Liquid Water will also decrease, in the same direction, since they are connected by a qualitative proportionality (P+), and with the same qualitative values of magnitude and derivative, since it has the correspondence Q between these two quantities.

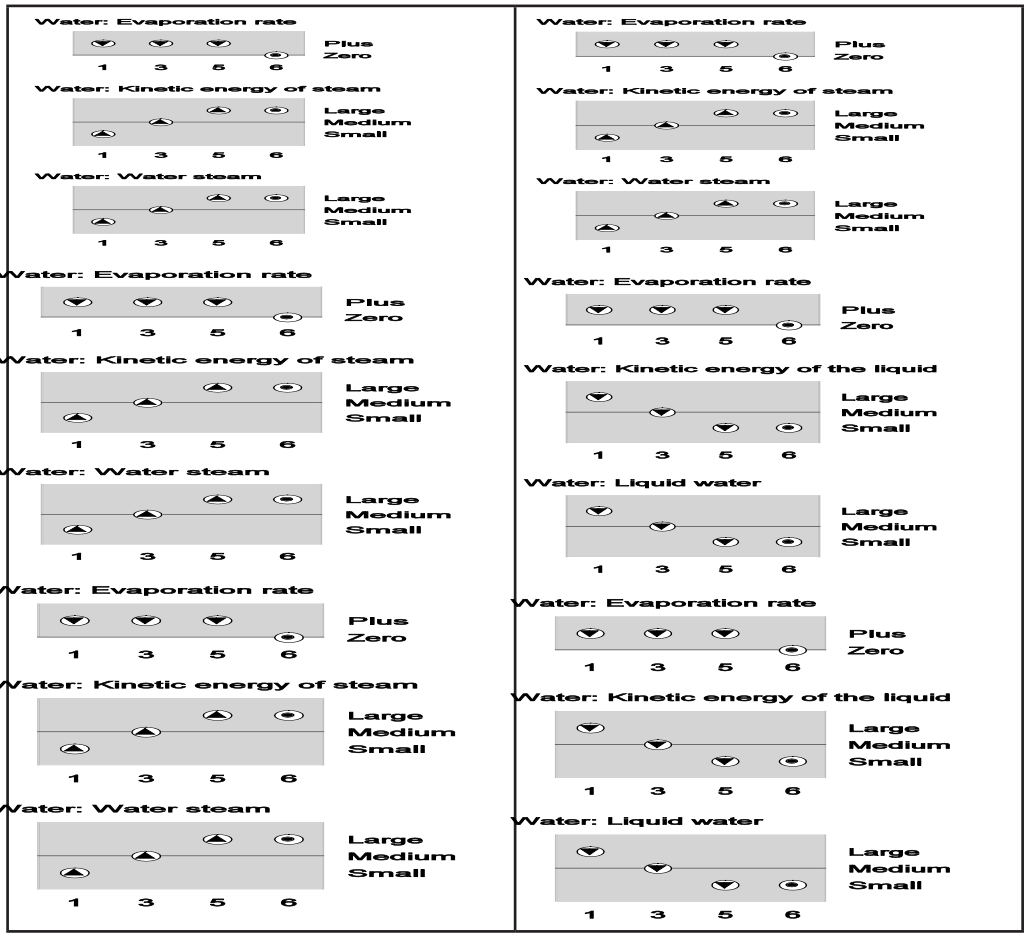
It is necessary to emphasize the meaning of the inverse correspondence (Q↓) between the KES and the KEL. In this case, Q↓ has two meanings: 1- the first, in relation to the very conceptual aspect of the modeling language, in which the values change simultaneously in an inverse way between the two quantities; 2- the second, in relation to specific knowledge of Sciences/Chemistry, in which it represents the transfer of energy from liquid to steam. When running the simulation of this model, the State Graph is generated, showing the change of states as in Figure 2.

Figure 2. State graph.



These are all the possibilities of change from that initial scenario. The values assumed by the quantities represented in the model during the simulation are shown in Value History Diagrams (VHD). In the simulation above, Figure 3 shows the selected trajectory, in which state 3 can continue changing to state 5, this one to state 6. Then VHDs allow us to understand the changes that occurred in the transitions between the states: [1] → [3] → [5] → [6]. This trajectory traces the behaviour of the system.

Figure 3. Value History Diagram (VHD).



The complete Simulation generates six states, one initial state and three final states. On the chosen trajectory [1,3, 5, 6], the Evaporation Rate has a Plus magnitude (positive), but a decreasing derivative, which remains in the states [1, 3, 5], until reaching 'Zero' and stable magnitude, in state 6. This rate has stabilized because it is controlled by the balancing feedback mechanism. The quantity influenced by the I+ of the rate, Kinetic energy of the steam, begins the trajectory at the 'Small' and increasing value, passing through state 3, with a 'Medium' value, and arriving at 'Large', at state 5, in which it stabilizes in the last state of this trajectory, [6]. The water vapor changes, during the trajectory, with the same values as its influencer, Kinetic energy of the vapor, as they are connected by the Q correspondence. The quantity influenced by the I- of said rate, starts the trajectory at the value 'Large' and decreasing, passing through state 3, with a 'Medium' value, and arriving at 'Small', in state 5, where it stabilizes in the last state of this trajectory, [6]. The same happens with liquid water, which is changed according to the same values as the kinetic energy of the liquid. In other words, the flow of energy turned liquid water into vapor and

the process came to a halt. The model does not include information about what could have happened later. Other examples of models involving the exchange of energy in the transition of physical states (for example, condensation, fusion, solidification) are found in Souza (2019).

Final Remarks

This work describes a qualitative model of the exchange of energy in the transition of physical states. Normally textbooks limit explanation to the transformation involved in changing the physical state of matter, especially with evaporation process, and they do not describe how these transformations happen. A QR approach has an added value because it focuses on the causal relations that determine the behaviour of the system. A description of the mechanism of change, the process of evaporation, indicates the origin of the dynamic phenomenon, which is then propagated to and observed in the rest of the system. QR has potential to impact other areas of science education not covered so far.

The work described here is part of a larger project that aims at the development of system thinking, on the part of students, through modeling. The use of qualitative models to support the system thinking by students is already being investigated and the results obtained so far are encouraging. Salles *et al*. (2006) found in an experiment that students exposed to qualitative models were consistent in the ability of recognizing objects and processes, building up causal chains and applying them to a given situation, assessing derivative values of quantities and making predictions about the consequences of changes in a system.

In relation to the model presented here, we had interesting results, and it was possible to relate, in a simple way, the dynamics of the transformation of the physical states of water. We represent energy exchanges, which are factors that actually promote the transition of physical states. Therefore, it is important to highlight that the model was initially created in a simpler way, but, in future research, we will present other concepts for real and more complex situations (for example, see Souza et al. 2019). The model, presented in this paper, allowed the causal relationships, the processes involved and the feedback mechanisms (in those made in LS4) to be properly explained. We hope that this model will be an instrument for future research and classroom work, both in Natural Sciences and Chemistry.

This paper signals enriching paths for a broader, more critical education, in which students are able to visualize and interpret operating relationships, phenomena of educational and scientific interest in a systemic way. In this context, qualitative modeling expands students' worldview, contextualizes problems, and enables student leadership in the learning process. This text leaves prospects for future work with another important contribution to the teaching of Science/Chemistry:

an educational material that makes it possible to discuss the transformation of the aggregation states of matter, involving energy exchanges.

Acknowledgement

The work described here was partially funded by Federal University of Uberlândia, Federal Institute Goiano and University of Brasilia.

References

Bredeweg, B. (1992). Expertise in Qualitative Prediction of Behaviour. Ph.D. thesis, University of Amsterdam, Amsterdam, The Netherlands, 1992.

Bredeweg, B. et al. (2013). DynaLearn – An Intelligent Learning Environment for Learning Conceptual Knowledge. *AI Magazine*, *34*(4), 46-65.

Forbus, K. D. (1984). Qualitative Process Theory. *Artificial Intelligence*, *24*(1), 85-168.

Mustapha, S.M.F.D. et al. (2002). Application of Qualitative Process Theory to qualitative simulation and analysis of inorganic chemical reaction. In: N. Agell & J. A. Ortega (Eds.) Proceedings of the International workshop on Qualitative Reasoning, (QR'02), pages 177-184, Sitges - Barcelona, Spain, June 10-12, 2002.

Oliveira, A. C. G. & Petrucci-Rosa, M. I. (2016). Recontextualizações e hibridismos em processos de elaboração e avaliação de Livros Didáticos de Química. Química Nova na Escola, 38(3), 273-283.

Salles, P. & Bredeweg, B. (2006). Modelling Population and Community Dynamics with Qualitative Reasoning. Ecological Modelling, 195(2), 114-128.

Souza, P. V. T et al. (2017). Um modelo de simulação baseado em raciocínio qualitativo para a educação em ciências. ACTIO: Docência em Ciências, 2(1), 162-183.

Souza, P. V. T. (2019). Modelos de simulação qualitativos como estratégia para o ensino de Ciências. Tese (Doutorado em Educação em Ciências) — Universidade de Brasília, Brasília.



Strand 2

Learning Science: Cognitive, Affective, and Social Aspects

Regina Soobard & Radu Bogdan Toma

Strand Chairs/Co-editors

Foreword

This chapter, titled Learning Science: Cognitive, Affective, and Social Aspects, is part of the ESERA 2023 Conference Proceedings and brings together eight significant studies that explore the multifaceted nature of science learning. Science education is not just about acquiring knowledge; it is deeply influenced by cognitive processes, emotional engagement, and social interactions. This chapter sheds light on how these interconnected aspects shape students' understanding and experiences in science education.

The cognitive dimension focuses on how learners process, comprehend, and apply scientific concepts. The studies in this chapter investigate diverse strategies to enhance cognitive engagement, from developing critical thinking skills to addressing misconceptions and promoting deeper conceptual understanding. These insights are crucial for designing instructional approaches that foster effective learning.

The affective aspect explores the emotions, attitudes, and motivations that influence students' science learning experiences. Understanding the emotional factors at play, such as students' confidence, interest, and anxiety levels, is vital for creating supportive learning environments that encourage positive engagement and persistence in science.

The social aspect highlights the role of collaboration, communication, and social interactions in the learning process. Science learning does not happen in isolation; it is enriched by peer interactions, group work, and discussions that help students co-construct knowledge and make sense of scientific concepts collectively.

The research presented in this chapter reflects a holistic approach to understanding science learning, recognizing that cognitive, affective, and social factors are deeply intertwined. By addressing how these elements interact, the studies offer valuable insights for educators, researchers, and practitioners aiming to enhance science education across diverse contexts.

We extend our sincere thanks to the authors and reviewers who contributed to this chapter. Their work not only advances our understanding of how students learn science but also provides practical implications for creating learning environments that are cognitively challenging, emotionally supportive, and socially engaging.

As you explore the proceedings in this chapter, we hope you gain valuable perspectives that will inspire further research and innovation in science education, ultimately contributing to more effective teaching practices and better learning outcomes for students.

Research Trends in Science Education from 2017 to 2021: A Content Analysis of Publications in ESERA Conference Proceedings

Georgios Ampatzidis[1], *Anastasia Armeni*[2]

[1]University of Thessaly, Greece

[2]University of Patras, Greece

Authoring academic publications is a fundamental task for researchers, serving both as a medium to introduce their work to the academic community and as a pivotal factor in their career advancement. Understanding research trends in these publications is essential for researchers, and systematic content analyses have been widely applied in various domains to gain insights into these trends. However, there has been a limited focus on analyzing the proceedings of science education conferences. This study's objective was to compile, organize, and offer a cogent summary of ESERA's research activity over the previous six years. We investigated a total of 639 articles from the ESERA conferences of 2017, 2019 and 2021, quantifying countries' research contributions, categorizing articles by topics and type, and assessing authorship patterns. We discuss relevant trends and emerging patterns.

Keywords: Science education, academic publications, content analysis

Introduction

Authoring publications stands as a major task for researchers (Tsai & Wen, 2005). On one hand, academic articles serve to introduce the work of researchers to the broader research community, and on the other hand, they play a vital role in advancing researchers' careers, influencing aspects like funding and recruitment into academic positions (Lee et al., 2009). Consequently, gaining insight into research trends, as mirrored in academic publications, holds great importance for researchers. A thorough examination of academic publications can provide valuable insights into the topics under investigation and the methodologies employed by the research community, effectively delineating the current state of a specific research field. In various research domains, systematic content analyses of articles published in relevant academic journals have been undertaken to serve this purpose, e.g. in mass communication (Kamhawi & Weaver, 2003), project management (Kloppenborg & Opfer, 2002), augmented reality applications (Chi et al., 2013), emerging markets (Kearney, 2012), tourism (Thomas et al., 2011). Within the field of science education, notable studies have been carried

out by Chang et al. (2010), Eybe & Schmidt (2001), Rennie (1998) and Tsai & Wen (2005). These studies involved an exploration of journals such as the International Journal of Science Education, Research in Science Education, International Journal of Science Teaching, Science Education, and Research in Science and Technological Education.

Although systematic investigation of the academic articles published in relevant academic journals has been carried out before, it seems that analyses of the proceedings of science education conferences are scarce. This study's objective was to compile, organize, and offer a cogent summary of ESERA's research activity over the previous six years. This systematic analysis of articles published in the ESERA proceedings of the previous three ESERA conferences (Dublin, 2017; Bologna, 2019; Braga [online] 2021) is expected to aid professionals working in the highly broad field of science education research in reflecting on historical trends. Our investigation concerns the country of origin of the authors, the frequency of the topic and type of published research, and the number of authors per article published. Thus, the research questions addressed here are the following:

1. How did authors from different countries contribute to the articles published in the selected ESERA proceedings (2017, 2019, 2021)?
2. How did the topics of the articles published in the ESERA proceedings varied across these three conferences (2017, 2019, 2021)?
3. What was the type of the articles published in the selected ESERA proceedings (2017, 2019, 2021)?
4. What was the number of authors of the articles published in the selected ESERA proceedings (2017, 2019, 2021)?

Methods

For this study, we investigated the 639 articles published in the proceedings of ESERA 2017, ESERA 2019 and ESERA 2021 conferences:

- Electronic Proceedings of the ESERA 2017 Conference. Research, Practice and Collaboration in Science Education (243 articles) (Finlayson et al., 2018).
- Electronic Proceedings of the ESERA 2019 Conference. The Beauty and Pleasure of Understanding: Engaging with Contemporary Challenges Through Science Education (238 articles) (Levrini & Tasquier, 2020).
- Fostering scientific citizenship in an uncertain world (Proceedings of ESERA 2021) (158 articles) (Carvalho et al., 2022).

For the first research question, each country's research contribution was quantified and ranked. Countries were given one point for every author mentioning them as their primary affiliation. For instance, if an article was authored by three authors from Germany, then Germany was given three points. Moreover, if an article was authored by three authors, one from Spain and two from Greece, then Spain was given one point and Greece was given two points respectively.

For the second research question, the analysis focused on the frequencies of articles falling into distinct topic categories, or strands, as outlined by ESERA. This involved a systematic calculation to provide insights into the distribution of research across various thematic strands.

For the third research question, the articles were categorized into empirical and theoretical. The category of empirical articles was consisted by the articles that present and discuss primary empirical data, while the category of theoretical works was consisted by articles such as position papers and literature reviews.

Finally, for the fourth research question, the frequency of articles authored by a specific number of authors was calculated, and then the data were categorized into relevant categories (articles authored by one, two, three, four, five, or more than five authors).

Results

The participants of ESERA 2017 and ESERA 2019 came from 39 different countries, while the participants of ESERA 2021 came from 34 different countries. In Table 1 the score of the top 20 countries for every conference is shown. The names of countries are the ones used for the affiliations of the respective authors. We noticed that some countries consistently ranked in the top 20 for all three conferences, such as Germany, Greece, Spain, and Brazil. Others, like Slovenia (only in ESERA 2019) and Sweden (in ESERA 2017 and 2019 but not in 2021), made appearances in the top 20 respectively.

Table 1. Country scores per conference.

	ESERA 2017		**ESERA 2019**		**ESERA 2021**	
Rank	**Country**	**Score**	**Country**	**Score**	**Country**	**Score**
1	Germany	125	Germany	112	Germany	104
2	Spain	67	Greece	95	Greece	66
3	Brazil	56	Spain	70	Spain	54
4	Ireland	42	Brazil	69	Brazil	45
5	USA	36	Japan	50	Italy	34
6	Greece	34	Italy	37	France	27
7	Switzerland	29	Austria	24	Japan	27
8	Sweden	26	France	19	Austria	25
9	UK	25	Ireland	19	Portugal	21
10	Japan	22	USA	19	USA	21
11	Italy	19	Sweden	14	UK	12
12	France	16	UK	13	Ireland	7
13	Austria	15	Turkey	11	South Africa	6
14	Finland	14	Switzerland	10	Canada	5
15	Croatia	13	Argentina	9	Chile	5
16	South Africa	13	Portugal	9	Argentina	4
17	Portugal	12	Slovenia	9	Norway	4
18	Denmark	9	Cyprus	8	Singapore	4
19	Canada	8	Israel	8	Turkey	4
20	Mexico	8	Canada	7	Czech Republic	3

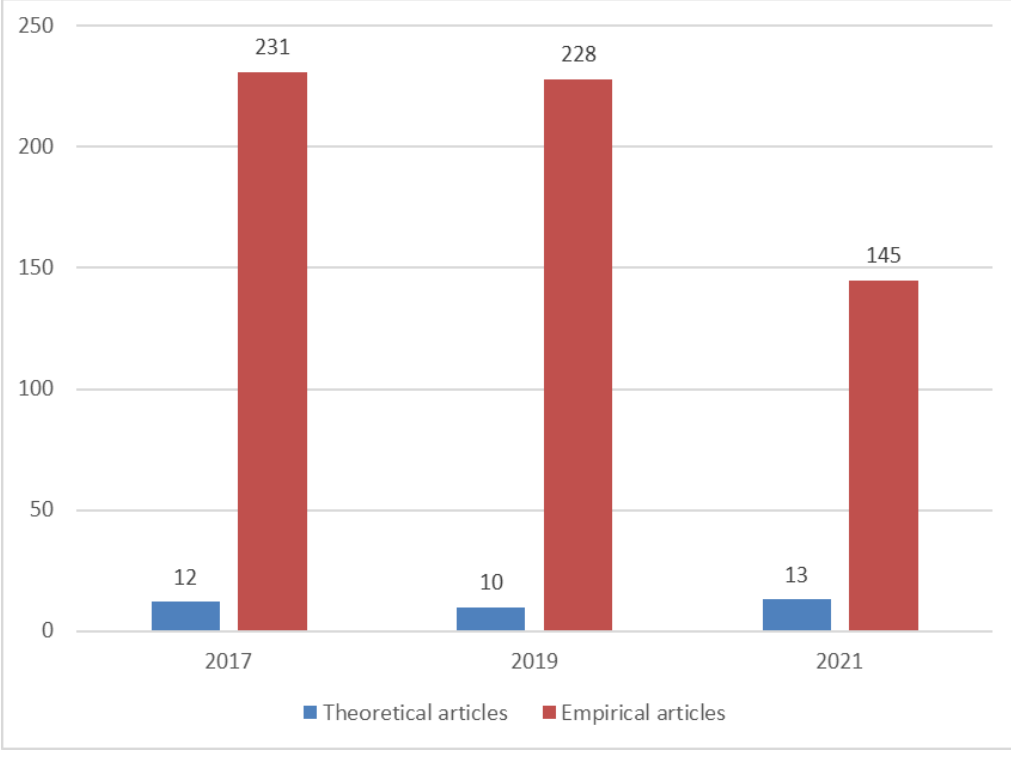
Concerning the article topics, we observed two distinct variations in categories among the conferences. Specifically, the category ‘Methodological Issues in Science Education Research’ appeared solely in the ESERA 2019 proceedings, while the category ‘Summer School Posters’ was exclusively featured in the ESERA 2017 proceedings. All other categories were consistent across the proceedings of the three conferences, facilitating meaningful comparisons. Notably, we identified a decline in the number of articles in the categories ‘Nature of science: history, philosophy, and sociology of science’ and ‘Learning science: Conceptual understanding’ over time (Table 2).

Table 2. Frequencies of topic categories per conference.

Topics	Frequencies		
	2017	2019	2021
Learning science: Conceptual understanding	24	17	10
Learning science: Cognitive, affective, and social aspects	21	17	11
Science teaching processes	12	14	8
Digital resources for science teaching and learning	14	22	10
Teaching-Learning Sequences as Innovations for Science Teaching and Learning	7	14	12
Nature of science: history, philosophy and sociology of science	17	7	2
Discourse and argumentation in science education	16	9	8
Scientific literacy and socio-scientific issues	11	11	13
Environmental, health and outdoor science education	16	12	18
Science curriculum and educational policy	6	4	3
Evaluation and assessment of student learning and development	15	15	6
Cultural, social and gender issues in science and technology education	11	7	9
Pre-service science teacher education	29	28	12
In-service science teacher education, continued professional development	11	31	18
Early years science education	9	8	3
Science in the primary school	8	8	4
Science teaching at the university level	8	9	11
Methodological Issues in Science Education Research	0	5	0
Summer School Posters	8	0	0

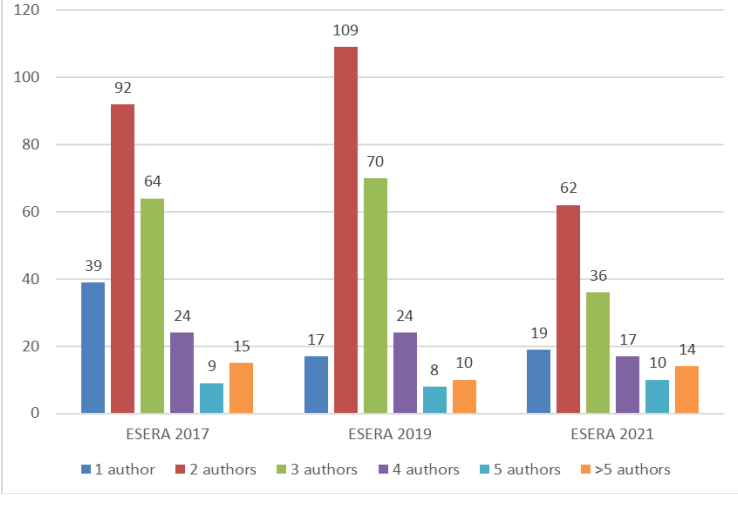
Concerning the type of the articles, we noted that the vast majority of published articles were empirical; only few articles fell into the theoretical articles category in all three conference proceedings volumes (Figure 1).

Figure 1. Frequency of articles by type.



Finally, the majority of articles were authored by up to three authors in all the three conference proceedings volumes. The articles authored by four or more authors were about 26% for ESERA 2021 proceedings and less than 20% for ESERA 2017 and ESERA 2019 proceedings as shown in Figure 2.

Figure 2. Frequency of articles by the number of authors.



Discussion

The ESERA 2021 conference was conducted online because of COVID-19 restrictive measures. However, this difference compared with ESERA 2017 and ESERA 2019 did not seem to have considerably influenced the nationalities of the participating authors. Some countries ranked at the top positions in all three conferences (e.g. Germany, Spain, Greece and Brazil). On the other hand, Ireland is a country that has lost ground in terms of representation after the ESERA 2017 that took place in Dublin and Italy is a country that gained ground in the two last conferences – it should be underlined that the ESERA 2019 took place in Bologna.

The fact that ESERA 2021 was conducted online resulted in substantially less articles published in the proceedings. This may seem paradoxical, as one may think that since ESERA 2021 was an online conference, it should attract more participants as they would not have to travel to a different country. However, this was not the case probably because research had also been restricted because of the measures against COVID-19 spread and there were less research results to be presented. Some of the differences we noted in the frequencies of articles topics are explained by the reduction of published articles in the ESERA 2021 proceedings. Nevertheless, interesting patterns that are not explained only by the reduction of published articles in the ESERA 2021 conference were noticed: for instance, the decline of interest in the topic 'nature of science' and the increase of interest in the topic 'teaching-learning sequences' from 2017 to 2021.

The vast majority of articles published in the proceedings of all the three conferences concerned empirical articles. The prevalence of empirical articles suggests that attendees are keen on presenting concrete findings and tangible evidence in the field of science education. Moreover, the majority of articles investigated were authored by up to three authors. This preference for smaller authorship teams could reflect a culture of collaborative yet closely-knit research communities, where scholars might engage in more personalized and focused investigations within science education.

In an effort to analyze the current state of research in science education as reflected in conference proceedings, we examined the articles published in three specific volumes of the ESERA conference proceedings. We focused on the ESERA conference as the subject of our study because it is very popular among science education scholars attracting a large number of participants from many different countries. Of course, the study of research trends in science education could be extended to explore older volumes of ESERA conference proceedings, or the proceedings of other popular science education conferences such as the NARST conference. Such an orientation could potentially prove interesting and valuable for future research.

References

Carvalho, G. S., Afonso, A. S., & Anastácio, Z. (Eds.). (2022). *Fostering scientific citizenship in an uncertain world (Proceedings of ESERA 2021)*. CIEC, University of Minho.

Chang, Y.-H., Chang, C.-Y., & Tseng, Y.-H. (2010). Trends of Science Education Research: An Automatic Content Analysis. *Journal of Science Education and Technology*, *19*(4), 315–331. https://doi.org/10.1007/s10956-009-9202-2

Chi, H.-L., Kang, S.-C., & Wang, X. (2013). Research trends and opportunities of augmented reality applications in architecture, engineering, and construction. *Automation in Construction*, *33*, 116–122. https://doi.org/10.1016/j.autcon.2012.12.017

Eybe, H., & Schmidt, H.-J. (2001). Quality criteria and exemplary papers in chemistry education research. *International Journal of Science Education*, *23*(2), 209–225. https://doi.org/10.1080/09500690118920

Finlayson, O. E., McLoughlin, E., Erduran, S., & Childs, P. (Eds.). (2018). *Electronic Proceedings of the ESERA 2017 Conference. Research, Practice and Collaboration in Science Education*. Dublin City University.

Kamhawi, R., & Weaver, D. (2003). Mass Communication Research Trends from 1980 to 1999. *Journalism & Mass Communication Quarterly*, *80*(1), 7–27. https://doi.org/10.1177/107769900308000102

Kearney, C. (2012). Emerging markets research: Trends, issues and future directions. *Emerging Markets Review*, *13*(2), 159–183. https://doi.org/10.1016/j.ememar.2012.01.003

Kloppenborg, T. J., & Opfer, W. A. (2002). The Current State of Project Management Research: Trends, Interpretations, and Predictions. *Project Management Journal*, *33*(2), 5–18. https://doi.org/10.1177/875697280203300203

Lee, M., Wu, Y., & Tsai, C. (2009). Research Trends in Science Education from 2003 to 2007: A content analysis of publications in selected journals. *International Journal of Science Education*, *31*(15), 1999–2020. https://doi.org/10.1080/09500690802314876

Levrini, O., & Tasquier, G. (Eds.). (2020). *Electronic Proceedings of the ESERA 2019 Conference. The Beauty and Pleasure of Understanding: Engaging With Contemporary Challenges Through Science Education*.

Rennie, L. J. (1998). Guest editorial: Improving the interpretation and reporting of quantitative research. *Journal of Research in Science Teaching*, *35*(3), 237–248. https://doi.org/10.1002/(SICI)1098-2736(199803)

Thomas, R., Shaw, G., & Page, S. J. (2011). Understanding small firms in tourism: A perspective on research trends and challenges. *Tourism Management*, *32*(5), 963–976. https://doi.org/10.1016/j.tourman.2011.02.003

Tsai, C., & Wen, M. L. (2005). Research and trends in science education from 1998 to 2002: A content analysis of publication in selected journals. *International Journal of Science Education*, *27*(1), 3–14. https://doi.org/10.1080/0950069042000243727

Use of Analogies in The Teaching of Science: Post-Festum and Heuristic Analogies

Nikolaos Fotou

School of Education, College of Arts, Social Sciences & Humanities, University of Lincoln, UK.

The paper switches attention to student self-generated analogies, as opposed to analogies generated by education researchers, teachers, or science textbook authors that the literature has largely focused on. It discusses the heuristic use and efficacy of student-generated analogies as a tool in approaching and correctly understanding what is new and unknown. Such use of analogies derives from research with a wide age range of students, who, when asked to make predictions in situations never considered before and to then provide explanations about these predictions, self-generated analogies. As it was found in their explanations, it was by reasoning on the basis of these self-generated analogies that some approached the situations they were presented with and made a scientifically compatible prediction for. This heuristic use of student self-generated analogies can be used in revealing a scientifically compatible way of reasoning on the basis of student self-generated analogies which can be also productively used in the process of science teaching.

Keywords: analogical reasoning, student-generated analogies, post-festum analogies.

Introduction

Science education research (e.g., Duit, 1991) has extensively argued for the use of analogies as an instructional tool that, when effectively used, can facilitate students' learning of abstract and, in many cases, difficult to understand concepts. The power of this instructional tool lies in that it allows students to form link what they know (the base domain which is the familiar situation) to what they are being taught and is unknown (the target domain which is the unfamiliar situation/concept being taught). Thus, when analogies are used effectively in the classroom, they can help students to harmoniously integrate the new concepts with their existing ones.

Whilst this paper does not contradict the findings of such prior research that has added instructional value to the use of analogies in facilitating new concepts learning, it shifts attention to the analogies students self-generate and their use in the science teaching. It focuses on the heuristic use of student-generated analogies as opposed to post-festum analogies (Wilbers & Duit, 2006) generated

by teachers for students who are asked to reason on their basis.

As the origin of the word heuristic denotes (adjective for the verb εὑρίσκω - heurískō in Greek, having the meaning of 'I discover/I find'), a heuristic analogy involves the construction analogical relations between a familiar and an unfamiliar situation. These analogical relations help students to 'find'/ 'discover' similarities between the two and draw inferences from the one to the other. Although similar is the function of student generated analogies more generally speaking in that it allows them to 'find'/'discover' similarities and understand (most often incorrectly) the unfamiliar on the basis of another more familiar, the term heuristic is used to differentiate between the analogies students generate and lead to an incorrect understanding (Fotou & Abrahams, 2016, 2017) and denote those that function heuristically in helping students to 'find'/ 'discover' relationships that are compatible with scientific conceptions and lead to a correct understanding of the target situation. This also differs to how heuristic analogies have been defined elsewhere in the literature (Wilbers & Duit, 2006).

The Study Purpose

Past research on analogies have mostly focused on their effectiveness as instruction devices in the teaching and learning of science and the degree to which analogies generated by teachers and/or researchers are then used by the students to transfer knowledge from the base domain they are provided with to the specific target domain being taught (e.g., Sarantopoulos & Chaparlis, 2004; Summers et al., 1997). The purpose has been to assess whether the similarities drawn from the base to the target domain are appositely made to lead to a scientifically correct inference from the one to another and thus a correct understanding. This past research has consistently found that, when appropriately utilised in the science classroom, analogies have the potential to facilitate the teaching and learning of science concepts.

Another facet of this science education research on analogical reasoning (e.g., May et al., 2006; Mozzer & Justi 2012; Pittman, 1999) has focused on analogies generated by students instead of these generated by someone else and them asked to use and reason on their basis. This has also added value to how student-generated analogies can be used in the science teaching in revealing the way they think but also the knowledge their analogical reasoning is founded upon (Fotou & Abrahams, 2023). Such findings on student analogical reasoning are suggestive of their use as diagnostic devices for either incorrect knowledge or a faulty reasoning that both can give rise to misconceptions.

The past work (e.g., May et al., 2006, Mozzer & Justi 2012) on student self-generated analogies, however, provides no insight, and thus adds no value to, student self-generated analogies that are compatible with the scientific perspective and a correct understanding. In other words, the extent to which

analogies students generate and are then used to draw and transfer conclusions between what is being presented to them (and may be unknown and unfamiliar like science concepts taught for the first time) in a harmonious and scientifically compatible way with what they may already be familiar with has not been explored. This a key area on analogical reasoning that this paper focuses on and has hitherto to be attended.

This paper presents parts of a larger project on student generated analogies and reasoning on their basis and discusses the heuristic use and efficacy of student-generated analogies as a tool in approaching and correctly understanding what is new and unknown. In other words, it focuses on the use of heuristic analogies and how these can be used in the process of facilitating conceptual understanding and therefore, as this paper argues, could be further developed, and encouraged by the teacher instruction.

Methodology

The findings on the use heuristic analogies derive from a mixed method cross-age study. The study sample consisted of students across the compulsory education levels with their age ranging from 10-17 years. Data were collected through the administration of a paper questionnaire followed by small group interviews. In the questionnaire, students were asked to make predictions in a number of situations they have not considered before -novel situations- and provide explanations about these predictions. In the group interviews, they were given the opportunity to further elaborate or provide clarifications to their predictions and explanations provided for these. More details about the study sample, and the research methodology, can be seen in previous work (e.g., Fotou, 2014, Fotou & Abrahams, 2023).

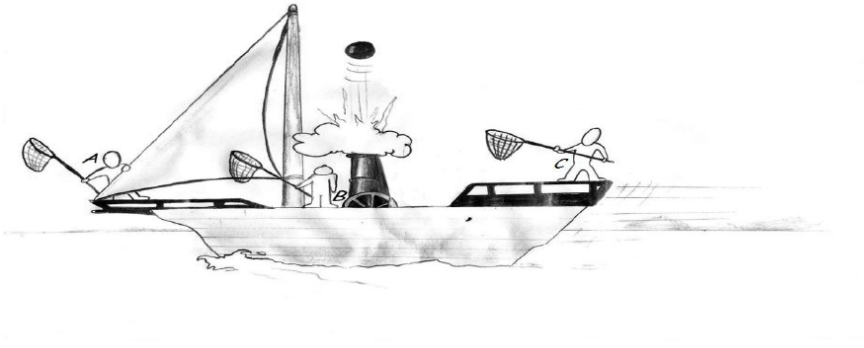
Findings

Below is one single case presented in this conference paper with one of the novel situations used in our study. This novel situation is similar with a situation in the form of a thought experiments in the dialogs of Galileo (Galilei, 1967). This was used to explain the relationship between motion and the reference frame.

In this novel situation, most of the students in the study sample reason analogically to make a prediction which they then expressed when explaining how they were led to this prediction. Although the majority were led to an incorrect prediction, there were instances of some whose analogical reasoning was heuristic in that the analogy they used and drew upon to make and transfer inferences to this novel situation, helped them reach conclusions and a prediction compatible with the scientific account.

Figure 1. A firing cannon in a moving boat

Novel situation 6: A firing cannon in a moving boat



While the ship keeps on moving at the same speed, the cannon fires a ball (as the figure shows). Which person is more likely to catch the ball?

A) Person A *B) Person B* *C) Person C*

Most of the students predicted that person it is person C who catches the ball. In the explanation of their predictions, they wrote about the idea of a straight down fall without taking into consideration the velocity that the boat possesses when the ball is fired. Hence, as the boat continues moving at that constant speed to the left, the ball falls, and lands behind the cannon. In coming up with this conclusion, students used different analogies based on their experiential knowledge of directly straight down motion than parabolic one. A ten-year-old explained, for example, that:

> Person C who is at the back of the boat will catch the ball. I have seen the same thing one day I went fishing with my father and some friends. I was walking with my friends and my father was following us behind. While we were walking, I tossed a small ball I was holding and it went backwards and my father caught it.

There were some instances, however, where students heuristically reasoned on the basis of self-generated analogies analogies which, as with the student above, helped them to draw inferences between the firing cannon in a moving boat and a situation they were more familiar with (the base domain), but were led to a correct understanding of the target situation and thus a prediction made compatible with the scientific account. A 17-year-old's response in this novel situation attest to this:

> I answered that person B is more luckily to catch the cannonball. I believe that as long as the boat continues moving at the same speed the cannonball does the same and this is why it falls at the same point from which it was fired. It the same with being in a bus and we toss a pen in the air. It lands in the same in the same spot, in our arms.

The answer shows that the student was able to draw on means of assessing the heuristic power of the analogy with conclusions drawn from and transferred to the novel situation (the target domain). The student also seems to be cognisant of some scientific principles (that the cannonball is fired with an initial horizontal velocity) that the use of the heuristic analogy either helped in either transferring to the target or helped in reaffirming this.

Conclusions

The findings show that the use of analogies in the teaching and learning of science could be extended from post-festum to heuristic analogies. This, on the one hand, and as discussed in other work (e.g., Pittman, 1999, Mozzer & Justi 2012), would provide science educators with a diagnostic form of assessment that can reveal when and how students draw inferences that are incompatible with the scientific account. It can be also used to diagnose he prior knowledge upon which student's reasoning and understanding is based and often gives rise to what has been termed in the literature as misconceptions (Fotou & Abrahams, 2023).

On the other hand, heuristic analogies generated by individual/groups of students, like those in the example presented in this paper, can be also used to introduce relevant concepts to the rest in the science classroom. Therefore, it is suggested here that student generated analogies that are heuristic can also help with the diagnosis and the identification of knowledge elements that are compatible with the scientific account and can be productively used in the teaching process.

References

Duit, R. (1991). On the role of analogies and metaphors in learning science. Science education, 75(6), 649-672.Fotou, N. (2017). Exploring the role of self-generated analogies in solving and reasoning about novel situations. In New Perspectives in Science Education Conference Proceedings. Filodiritto Publisher.

Fotou, N. (2014). Students' predictions in novel situations and the role of self-generated analogies in their reasoning (Doctoral dissertation, University of Leeds).

Fotou, N., & Abrahams, I. (2023). From the known to the unknown: The role of spontaneous and self-generated analogies in students' predictions about novel situations. Research in Science & Technological Education, 41(3), 946-960.

Fotou, N., & Abrahams, I. (2016). Students' reasoning in making predictions about novel situations: the role of self-generated analogies. In *Insights from Research in Science Teaching and Learning: Selected Papers from the ESERA 2013 Conference* (pp. 123-138). Springer International Publishing.

Galilei, G. (1967). Dialogue concerning the two chief world systems (2nd ed., ed. and trans Stillman Drake). Berkeley: University of California Press.

May, D. B., Hammer, D., & Roy, P. (2006). Children's analogical reasoning in a third-grade science discussion. Science Education, 90(2), 316-330.

Mozzer, N. B., & Justi, R. (2012). Students' pre-and post-teaching analogical reasoning when they draw their analogies. *International Journal of Science Education*, *34*(3), 429-458.

Pittman, K. M. (1999). Student-generated analogies: Another way of knowing?. Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching, 36(1), 1-22.

Sarantopoulos, P., & Tsaparlis, G. (2004). Analogies in chemistry teaching as a means of attainment of cognitive and affective objectives: A longitudinal study in a naturalistic setting, using analogies with a strong social content. *Chemistry Education Research and Practice*, 5(1), 33-50.

Summers, M., Kruger, C., & Mant, J. (1997). *Teaching Electricity Effectively: A Research-Based Guide for Primary Science*. Association for Science Education, College Lane, Hatfield, Herts AL10 9AA, England, United Kingdom.

Wilbers, J., & Duit, R. (2006). Post-festum and heuristic analogies. *Metaphor and analogy in science education*, 37-49.

Meeting Between Spinoza and Gilles Deleuze: From Affection to Learning in Sciences

Amanda Jardim, Carolina Souza

Federal University of São Carlos, Brazil

This paper is part of an ongoing master's research project. It aims to propose Spinoza's concept of affection and Gilles Deleuze's notion of "learning", based on the experience of encounter with signs, as an alternative to think about Science Education. The article questions Science Teaching based on Modern Science, traditionally focused on reason, and on a notion of objectivity, in the proposal of a Science Education with all, and not for all.

Keywords: Affection, science education, signs.

Introduction

Laws, theories, hypotheses, laboratories, formulas, reactions, investigations. Words, concepts and signs that cross Science Teaching are immeasurable, they go beyond scientific representations. But how are they thought of in research and articulated in school practice? The remnants of traditional pedagogy still cross Education as a whole, so that the school is still seen as the place of "transmission of knowledge accumulated in history", the teacher "as central figure" the transmitter of knowledge, and the student is *alumini, that* is, without light, and his duty would be to assimilate knowledge, memorize and do the tasks, in a passive posture. (Oliveira, 2009). Therefore, it is not difficult to imagine how these signs of Science Teaching permeate the science classes.

The word teaching refers to the sign, and when we opt for a teaching process that does not start from lived processes, teaching becomes only fixing signs, giving the idea that the student bodies are voids. A void to be filled, where the teachings will be deposited, ignoring the idea that "one does not know how a body learns", as proposed by Deleuze. In this way, the act of teaching is nothing more than an act of imposing signs and commands (Deleuze & Guattari, 1980/1995).

> The teacher doesn't question herself when she questions a student, just as she doesn't question herself when she teaches a grammar or calculus rule. She "teaches", gives orders, commands. [...] The machine of compulsory education does not communicate information, but imposes semiotic coordinates on the child with all the dual bases of grammar (masculine-feminine, singular-plural, noun-verb, subject of enunciation-subject of enunciation, etc). The elementary unit of language - is the

watchword (Deleuze & Guattari, 1997, p. 11-12).

Given these perspectives, the present article aims to problematize the idea of a transition from science teaching to science education. In this sense, Lima (2022), in his text entitled "Você ensina ou você educa?" (Do you teach or do you educate?), invites us to think about the processes of teaching and educating. Starting from the etymology of the words, in which education comes from the Latin and refers to the idea of leading out; to develop the potential that is within each one. The word education, therefore, brings with it a concept of human nature. We are not a tabula rasa. "We are not a blank slate on which knowledge can be deposited". Science Education is a broad approach. It involves thinking about society, public policies, diversities, genders, social commitment and not only instruct, train, charge lists of exercises. Even because, according to Reiss (2005), the success of Science Education depends on the meaning and personal value that students attribute to what is being taught.

In the contribution of the thought of a displacement from Science Teaching to Science Education, we also propose to think of the body in its singularity, facing the act of learning. As a construction that happens from the inside out. Getting away from the traditional conception of Science Teaching, which says to deposit the light, in the being without light. In this way, the notion of "learning" which we will discuss in this article does not have the sense of acquiring knowledge or a passage from not knowing to knowing, but of experience and meaning, according to Gilles Deleuze. Thus, it is based on a different ethics and aesthetics, it focuses on practice and distances itself from the expository lessons in which the teacher is the centre and the students are passive. In this way, the Philosophy of Difference proposed by Deleuze, which emphasizes creation in the teaching-learning process, with emphasis on aesthetics, ethics and politics, within another ontology, is a potent reference for this article.

For a learning that goes through the bodies in their singularity, we will focus in this article on the theoretical contributions on the joyful affections of Spinoza. The objective of this discussion is to present the importance of a Science Education that goes through the body. The theoretical discussion is supported by Nietzsche's thought, that the body is constituted by a relation of forces in conflict. Also, in the idea that thought is creation and not mere reproduction presented by Deleuze in many sources, to support the notion of learning.

Theorical Contributions

In Deleuzian thought, from his reading of Nietzsche, the body can be thought as an encounter of forces. The body is always a relation, which puts this body in movement, with different speeds; The body is always a relationship, which places it at rest and in movement, at different speeds. Bodies are modes, complex relations of speed and slowness (Deleuze, 1970/2002). Therefore, within this

thinking, the body escapes the idea of a totalized body, organized and defined in forms and functions, the body is not, it is being, varying from body to body, differentiating itself from the others and especially from itself, according to the dynamics of rest and movement, speed and slowness, as well as its ability to affect and be affected. Therefore, thinking about the encounters and ways of life that enhance a body, is also part of thinking about education.

In Deleuze's concept of learning, learning refers to the "encounter with signs". The signs are emissions that can be projected by materials, objects, beings, people, "everything is a sign" and "there are only signs" (Deleuze, 1984/2010). The encounter with the signs may affect, and when crossed by a sign, the body differentiates itself, lives a problematization experience, invents problems that causes the need to think, interpret, forces the body to move to seek solution and meaning to the emitted sign (Kastrup, 2001). "To interpret is to give meaning, to impose an order, a form, a direction, is to give a sign to the shapeless and chaotic mass of things in the world. To interpret is not to reveal, discover, identify, but to create, invent, produce" (Corazza & Tadeu, 2003).

> Learning takes place in the search for a gestural response to a problem generated by the encounter between two bodies, the swimmer and the water. This response is only given by the singularity of each body, and therefore varies from body to body. This is why "[...] it is so difficult to say how someone learns". (Deleuze, 2006, p. 48).

The present investigation, it is understood that what happens in the science classroom, incorporating the ethical, emotional, cultural and social dimension, is not limited to the appropriation of a content or a technique. What matters are experiences, understood as what passes through us, what happens to us, what touches us. This experience takes place in the body and through the body, this experience is different, it doesn't have a pre-defined goal, as Raul Seixas, a successful Brazilian musician in the 1970s, sings in his song "Metamorfose Ambulante" (Walking Metamorphosis), released in 1973, "it's boring to reach a goal in an instant", this is an experience that happens, it's the surprise of an encounter that happens in an instant, and this instant escapes the time of the pointers of a clock, it lasts as long as it has to. It's an experience that can't be dominated, manipulated or remade, it's an unconditional and unconscious surrender, it's a becoming.

Deleuze (2010) explains that in Proust's work "La recherche du temps perdu", lost time refers to the time that has passed, but not only that: It is "past time"; it is also the time that is lost, as in the expression "wasting time". The dimensions of time take into account time lost, time found and time to be found. And learning takes place, from this perspective of encounter, by letting go of obligations and reason, allowing oneself to be involved. So, in this way of thinking, educational

practice escapes from a movement that aims to develop in favor of seeking to involve.

Learning essentially concerns signs, in a personal and temporally located and situated experience. The learning of signs corresponds to "living the experience, from inside" (Deleuze, 2010), which goes beyond the stated aims and principles. Learning depends on being sensitive to the sign: "One only becomes a carpenter by becoming sensitive to the signs of wood, and a doctor by becoming sensitive to the signs of illness" (Deleuze, 1964/2010). And this sensitivity refers to: "to consider the world as something to be deciphered is undoubtedly a gift. But this gift would run the risk of remaining hidden in ourselves if we didn't have the necessary encounters" (Deleuze, 2010).

In this sense, the process of deciphering and interpreting the signs goes through the body, a world of its own, which, at random encounters, creates a meaning for itself. Therefore, the problems experienced are singular, the interpretations are multiple, because the sign is not universal. This movement is given by the need and desire that activates the forces of the bodies in their singularity. Since, "A creator only does what he or she absolutely needs to do" (Deleuze, 1987).

However, learning with signs doesn't just involve the chance of encountering and being sensitive to signs, it also forces us to think, because when a sign appears "we are sure that it hits us from the outside, but we still don't know what it means. It has the force of a question that forces us to think, of a problem that demands a solution" (KASTRUP, 2001, p. 20). This force acts as violence on thought, which is established as a necessity. However, this need "is not defined negatively, it is not a lack, but it is positive. It is a need that expresses the opening of a problem, not the impossibility of a solution". (Kastrup, 2001).

Lorenzetti and Delizoicov (2001) compare Science Education to the reading and writing process that bodies make of their social context. For the authors, Science Education must be able to promote the individual recognition of the symbols and codes of science. And therefore Science Education takes place from a movement of creation, ways of giving meaning to the signs of science within an own universe. (as cited in Brito & Ramos, 2016).

Between Signs and Affections: A Possibility of Science Education

For Spinoza (2007), our bodies are affected by encounters, which happen at random and can be good or bad. So, in a good encounter, the bodies are composed, there is a joyful affection, the body increases its power to act. And in a bad encounter, the opposite, the body breaks down, there is a sad affection that decreases the power to act. In this way, the encounter is the guiding thread of affections. It is this fortuitous encounter between bodies that makes affection happen, and once the body is affected, it experiences something else, which drives it towards creation, which makes another way of existing possible.

> Affect is not reduced to an intellectual comparison of ideas; affect is constituted by the lived transition or the lived passage from one degree of perfection to another, while this passage is determined by ideas. But in itself, it does not consist of an idea, it constitutes the affect (Deleuze, 2019, p. 41).

It is for this reason that we cannot conceive of the idea of a totalized body, because we understand affect as what composes bodies, so that, in this line of thought, bodies would be the result of the constant relationships that are created through affect, given that: "affect is what happens to us concomitantly with our contact with the world". (Beccari & Bedore, 2017). Affection is also the event that makes us adhere to the world. Through it we are propelled into a movement of creation, and when the body creates, it differentiates itself, creates another form of existence.

An encounter with a sign that affects us promotes more than affections, which increase or decrease the power to act. Besides this passage from one state to another, there is also a mark, the trace of the action of a body on the other. Immersed in these concepts Yonezawa (2015) reflects: "a body only begins to know another body by its affections, that is, by how it is marked. In other words, in the body, the power to know and to think, that is, to have an idea, is only activated when there is affection". Therefore, to promote good encounters is to launch the body to know, a dive in the world of matter, is to experiment, is to create, is to promote differentiations.

> Knowledge is happiness. Spinoza's ethics finds its actuality here, because it is, in its own condition, a system that claims affectivity and seeks the happiness that comes from it, under the sign of a feeling that comes from our totally appropriate power to be in act. The being in act is an experiment, a difference, the unexpected, that which is yet to come (Lins, 2008, p. 49).

In this line of thought, the science teacher, besides being an emitter of signs of nature, is also an attractor of affections, so that: "he attracts to the matter and not to a ready-made knowledge. He/she is someone who performs the function of leading the process, the expedition to an unknown world, of making contact happen". (Kastrup, 2001). But for this movement to be possible, as well as teachers being sensitive to the signs of science, they must be sensitive to their students, knowing their students capacity to affect and be affected.

This allows them to escape from educational practices that promote sad affections, to abandon the content-dump, the given interpretations, the scripted experiments, the lined-up desks, the sepulchral silence, the passive student. To get involved, to share joyful classes/encounters, in which affections are triggered, bodies sing, dance, draw, act, create experiments and interpretations, re-exist. A space for

collective exchange and enrichment, as Spinoza also points out:

> If someone has done something that he imagines affects others with happiness, he also will be affected with happiness, which will be accompanied by the idea of himself as the cause, in other words, he will regard himself with happiness. If, on the other hand, he has done something that he imagines affects others with sadness, he will regard himself as sadness (Spinoza, 2007, p. 199).

It is in this sense that we meet and are affected by Spinoza's philosophy. We see it as a possibility for a joyful science education. Science classes would then be a territory of joyful encounters with the signs of nature, they would be classes/encounters, in which signs are triggered, enabling bodies to affect, create and affirm life. Boosting thinking beyond the process of rationalization, but towards a movement of creation that brings bodies closer to the world and to themselves.

References

Bedore, R. C. & Beccari, M. N. (2017) *Aisthesis: uma breve introdução à estética dos afetos* [Aisthesis: a brief introduction to the aesthetics of affections]. Gearte Journal, v. 4, n. 3. doi: 10.22456/2357-9854.74040. Retrieved from https://seer.ufrgs.br/index.php/gearte/article/view/74040

Brito, M. R. & Ramos, M. N. (2016). *Inspiração deleuziana para pensar a educação em ciências* [Deleuzian inspiration to think science education]. Dialogia, 23, 137-146. doi:10.5585/Dialogia n.23.5626

Corazza, S. & Tadeu, T. (2003). *Composições* [Compositions]. Belo Horizonte: Autêntica.

Deleuze, G. (2010). *Proust e os signos* [Proust and the signs]. Rio de Janeiro: Forense. (Original work published 1964)

Deleuze, G. & Guattari, F. (1995). *Mil Platôs: Capitalismo e esquizofrenia (vol.2)* [A thousand plateaus: Capitalism and schizophrenia]. São Paulo: Editora 34. (Original work published 1980)

Deleuze, G. (2002). *Espinosa: filosofia prática* [Espinosa: Practical philosophy]. São Paulo: Escuta. (Original work published 1970)

Deleuze, G. (2019). *Cursos sobre Spinoza* [Courses on Spinoza]. Fortaleza: Eduece.

Deleuze, G. (2006). *Diferença e repetição* [Difference and repetition]. São Paulo: Graal.

Deleuze, G. (1987). *What is a creative act*. In Tuesday lecture series, Femis film foundation. Paris.

Kastrup, V. (2001). *Aprendizagem, arte e invenção* [Learning, art and invention]. Psicologia em estudo, 6, 17-27. Retrieved from https://www.ufjf.br/ebrapem2015/files/2015/10/gd7_marta_oliveira.pdf

Lima, N.W. (2022). *Como estruturar uma aula que faça sentido?* [How to structure a lesson that makes sense?]. Porto Alegre: Author's Edition. Retrieved from https://sites.google.com/view/nathanwlima/materiais-para-download?pli=1

Lins, D. (2008). *A alegria como força revolucionária* [Joy as a revolutionary force]. Fazendo rizoma: pensamentos contemporâneos; São Paulo: Hedra.

Lorenzetti, L., Delizoicov, D. (2001). *Alfabetização científica no contexto das séries iniciais* [Scientific literacy in the context of early grades]. Ensaio Pesquisa em Educação em Ciências, 3, 45-61. doi: https://doi.org/10.1590/1983-21172001030104

Reiss, M.J. (2005). *The Importance of Affect in Science Education*. Science & Technology Education Library, 29, 17-25. Doi: https://doi.org/10.1007/1-4020-3808-9_2

Oliveira, R. M. A. (2009). *Ensino e aprendizagem escolar: algumas origens das ideias educacionais* [School teaching and learning: some origins of educational ideas]. São Carlos: Edufscar.

Spinoza, B. (2007). Ética [Ethic]. Belo Horizonte: Autêntica Editora. (Original work published 1677)

Yonezawa, F. H. (2015). *Só a alegria produz conhecimento: corpo, afeto e aprendizagem ética na leitura deleuzeana de Spinoza* [Only joy produces knowledge: body, affection and ethical learning in Deleuzian reading of Spinoza]. Educação Teoria Prática*,* 25, 186-199. Doi: https://doi.org/10.18675/1981-8106.vol25.n48.p186-199.

The Role of Open Schooling in Community Efforts to Tackle the Silent Pandemic of Antimicrobial Resistance

Irene Drymiotou, Constantinos P. Constantinou

University of Cyprus, Cyprus

This case study reports on an open schooling initiative for enhancing students' interest in science and understandings of relevant careers in the context of a public health issue. Framed within the design-based research paradigm, this pilot study enacted an open-schooling Teaching-Learning Sequence around the socio-scientific issue of Antimicrobial Resistance according to which students engage in scientific practices and bidirectionally collaborate with different stakeholders in the wider community. The participants were 74 students (13-14 years old) attending an urban secondary school. Data were collected utilizing a questionnaire and interviews with the students before and after the intervention. The findings from the pilot study suggest that students' genuine engagement in scientific practices, using actual equipment and collaborating with experts in out-of-school settings, while also promoting and sharing knowledge within a group setting, stimulated their interest and contributed to an improved perception of science in practice and, consequently, science-related careers. The findings may inform the design of open-schooling teaching and learning activities and provide practical recommendations for curriculum design and classroom practices with the potential to enhance students' interest in science and relevant careers.

Keywords: Interest, open schooling, antimicrobial resistance.

Introduction

Globally, over the last two decades, societies face complex societal issues such as public health threats and climate change implications, that urgently call for citizens to act toward achieving sustainability (UNESCO, 2020). Recent data published by the Organisation for Economic Co-operation and Development [OECD] (2021), indicate that not only cognitive skills but also transversal skills and attitudes acquired through lifelong learning will help individuals to thrive in a post-truth era. Tackling these growing issues demands scientifically literate citizens who can engage in public discourse and evidence-based reasoning and take informed decisions for a better quality of life. This set of skills can be fostered in science learning by practicing scientific thinking (Hazelkorn et al., 2015). Nevertheless, students' declining interest in science remains an issue of concern worldwide which also has a domino effect on the percentage of STEM

education graduates (26%) particularly in Europe (OECD, 2019). Previous work (Drymiotou et al., 2021a) has shown that opportunities for active engagement in scientific practices with experts, connections between STEM careers and curricular topics as well as transferring science concepts to a real-life personally relevant context may succeed in enhancing students' interest in science and in studying and pursuing STEM careers. Hence, an important mission of educational institutions is to provide such opportunities to students and nurture them into becoming responsible citizens.

In doing so, the present pilot study aims to investigate opportunities to enhance students' interest in science by promoting the mechanism of *Open Schooling*. Open schooling is conceptualised within the context of project MULTIPLIERS (https://multipliers-project.org/), as an educational perspective in which schools become open to society by bidirectionally collaborating with different stakeholders to (a) improve community well-being by raising awareness and co-creating solutions to both personal and socially relevant problems; (b) engage in inquiry processes, construction of knowledge, creative action and dissemination processes at a local and global level, and (c) enrich the curricula and pedagogical approaches in schools while promoting meaningful learning and competence development (Constantinou & Papadouris, 2012). This conceptualisation derives from a systematic review of existing good practices (i.e., a set of EU Open Schooling Calls, 12 EU-funded projects, 9 initiatives in the partner countries, and 50 articles) and a needs analysis using focus group interviews with 45 stakeholders.

This theoretically and empirically rooted conceptualisation guided the formulation of the framework for developing and enacting an open-schooling Teaching-Learning Sequence (TLS) (Papadouris & Constantinou, 2016, 2017) as shown on Table 1. Hence, this study aims to investigate whether open-schooling educational actions can enhance students' interest in science and their understandings of science careers by responding to the following research question:

Do open schooling educational actions influence:

a) students' interest in science?

b) students' career awareness?

Methods and Context

This study embraced a case-study approach with a case being defined as a group of 74 students (13-14 years old, 42 identified as female and 32 identified as male students) (Merriam & Tisdell, 2015) to allow a deeper insight into the ways in which open schooling educational actions influence students' interest in science and career awareness. The participants were divided into four

classrooms attending an urban secondary school in Cyprus taught by two science teachers.

Framed within the 'Design-Based Research' (DBR) paradigm (Brown, 1992) the research is design-driven and intervention-focused. The setting of this study interplays within the classroom environment and the wider community in collaboration between researchers, teachers, STEM experts, and civil society organisations. The study unfolded based on the open-schooling TLS framework focusing on the socio-scientific issue of Antimicrobial Resistance (AMR) in the teaching unit of 'Microbes and Disease' (Drymiotou & Constantinou, 2023). AMR is a critical issue around the globe affecting our health and putting our lives at risk. Overwhelming evidence shows that increased use of antibiotics, over-prescription, and overconsumption are leading to an increase in resistant bugs (European Centre for Disease Prevention and Control [ECDC], 2022). Cyprus where this study takes place ranks amongst the EU/EEA countries with the highest consumption of antibiotics (ECDC, 2022). A vital step towards tackling this issue is education and informed action. Table 1 describes the ongoing TLS that was co-developed between researchers, teachers and biology experts. The enactment consisted of 20 sessions (1 session= 50') excluding after-school hours.

Table 1. Open schooling Teaching and Learning Sequence around AMR

Open Schooling TLS Phases	**No of Sessions (S) and Short Description**
Identification and Exploration of the SSI of relevance to the local community (5 sessions)	**S1**. Introduction to AMR building on previous knowledge. **S2-3**. Students explore the issue of AMR in stakeholder groups (i.e., pharmacists, policymakers, farmers, patients, academia, media) using authentic media and evaluate the information using the TAAP test (Time-Authority-Accuracy-Purpose test, adjusted version of CRAAP test). **S4**: Introduction to argumentation: identifying and formulating arguments. **S5**: Debate within the stakeholder groups on the topic 'We should ban the use of antibiotics for the treatment of flu. Do you agree or disagree?'
Engagement in the Open Schooling Science-learning project (10 sessions)	**S6**. Experts visit the school, discuss the importance of AMR, and share an overview of the visit to the premises the day after. **S7-9**. Students visit experts' premises and engage in authentic activities and scientific practices. **S10**. Experts visit the school to summarise the key messages from S7 and show photos of bacteria culture collected from students' hands. **S11**. Presentation of students' mission to start an awareness-raising campaign for AMR consisting of different projects. **S12-15**: Development of the Open Schooling projects.

Acting as knowledge multipliers (5 sessions)	**S16**. Presentation of the projects to the local science fair (sCYence fair). **S18-20**. Presentation of the projects to the awareness-raising campaign at the central square of the capital.

Data Collection

To investigate whether open schooling influences students' interest and career awareness, a combination of quantitative and qualitative data was collected. Quantitative data were collected using an adjusted version of the Scenario Evaluation with Relevance and Interest (SERI) (Kang et al., 2021). Qualitative data were retrieved from interviews with the students on the added value of their experience with the open schooling TLS, in terms of enhancing their interest in science and raising awareness about science careers. Interviews with the students were conducted after S11 and Phase 3.

The adjusted version of the SERI questionnaire consists of 16 items related to interest (8 items), self-efficacy (4 items) and career awareness (4 items) on a 4-point Likert scale (from *strongly disagree* to *strongly agree; not well informed* to *very well informed*). For this pilot study, we considered the items related to interest and career awareness as displayed in Image 1 below. The questionnaire was administered before the beginning of Phase 1 and after Phase 3.

Figure 1. Snapshot of the questionnaire given as a pre - and post-test.

SCIENCE INTEREST (8 items)				
How strongly do you agree?	Strongly disagree	Disagree	Agree	Strongly agree
I really like science.	☐	☐	☐	☐
Doing science is one of my favourite activities.	☐	☐	☐	☐
I enjoy science courses.	☐	☐	☐	☐
I would like to find out much more about some of the things we deal with in our science class.	☐	☐	☐	☐
Science is relevant to me.	☐	☐	☐	☐
Science is important for helping us to understand the natural world.	☐	☐	☐	☐
Science is valuable to society.	☐	☐	☐	☐
I find that science helps me to understand the things around me.	☐	☐	☐	☐
SCIENCE CAREER AWARENESS (4 items)				
How informed are you?	Not well informed	Somewhat informed	Informed	Very well informed
Science-related careers that are available in the job market.	☐	☐	☐	☐
Where to find information about science-related careers.	☐	☐	☐	☐
The steps students need to take if they want a science-related career.	☐	☐	☐	☐
Employers or companies that hire people to work in science-related careers.	☐	☐	☐	☐

Qualitative data were also collected through ten focus group interviews (8-10')

with 20 students (4 in each group) after S11 and Phase 3, and two science teachers (in total, twelve interviews). The interviews were semi-structured in both cases and focused on providing insights into the ways in which the open schooling works (or not) towards enhancing interest in science and career awareness (e.g., successful activities/methods, obstacles, suggestions for improvement). The students were selected based on their scores on the questionnaire including a diversity of levels concerning interest in science respecting gender diversity as well (11 female and 9 male students).

Data Analysis

Quantitative data were analysed using a pre- and post- data comparison to provide an overall indication of students' interest in science and awareness about science careers using the SPSS software (version 29). Data were skewed for one of the two variables, and a Wilcoxon Signed Ranks Test was run to find the difference between the two measurements also comparing the median value of the changes in students' answers before and after the intervention for each of the studied variables. Qualitative data from the interviews with the students were analysed using open coding with respect to the features that seemed to enhance interest and career awareness.

Results and Discussion

Findings from the Quantitative Analysis

Overall, the findings from the quantitative analysis indicated that the intervention had a positive impact on both students' interest in science as well as their awareness about science-related careers. According to the results (Table 2), the post-test scores for students' interest in science were statistically significantly higher than the pre-test scores, Z= -3.58, p<.001. Similarly, the post-test scores for students' career awareness in science were statistically significantly higher than the pre-test scores, Z= -3.25, p=.001.

Table 2. Comparing students' science interest and career awareness before and after the intervention using the Wilcoxon Signed Ranks Test results.

Variable (post-pre)	Ranks	N	Mean Rank	Sum of Ranks	Z	p
Interest	Neg.	14	33.43	468.00	-3.58	<.001
	Pos.	48	30.94	1485.00		
	Ties	12				
Career awareness	Neg.	19	29.32	557.00	-3.25	.001
	Pos.	45	33.84	1523.00		
	Ties	10				

According to further analysis in each of the question items it was found that students' post score for question item 2 'Doing science is one of my favourite activities' had a statistically significant difference than the pre-test score compared to the other items, Z= -6.96, p<.001. Likewise, considering the career awareness-related items, statistically significant differences were found in 3 out the 4 question-items (i.e., 'Science-related careers that are available in the job market'; 'The steps students need to take if they want a science-related science'; and 'Employers or companies that hire people to work in science-related carers'). In these three question-items, the students' post-test scores were statistically significantly higher than the pre-test scores, *Z*= -2.12, p<.05, *Z*= -3.46, p<.001, and *Z*= -3.11, p<.002, respectively.

Findings from the Qualitative analysis

The qualitative analysis using open coding with respect to the features that seemed to promote interest in science facilitated the emergence of three categories with their frequencies as shown on Table 3 below. Specifically, the categories shed light on what the students found more interesting/liked more which provides further insights on the ways in which the open schooling mechanism influenced their interest in science. The frequencies indicate the number of times the students mentioned these features in each category and code.

Table 3. Features that enhance students' interest in science according to the analysis of the interviews with students (n=20).

Features	Frequencies
Novelty	26
Open schooling events	10
Interaction with experts in an authentic environment	9
Scientific practices	7
Knowledge	16
Sharing knowledge	9
Learning/knowledge gained (AMR and careers)	7
Social relatedness	5
Working in groups	3
Social interaction	2

As can be derived from Table 3, the novelty aspect (n=26) was found to have influenced students' interest significantly. Students' participation in the open schooling events was perceived as remarkably interesting for the students (n=10) as well as their interaction with the experts (n=9) and the engagement with the scientific practices (n=7). The excerpts on Table 4 below from the interviews

with the students are indicative of the supporting evidence.

Table 4. Excerpts from interviews with the students indicate novelty as a feature that enhances interest in science.

Category	Code	Excerpts from the interviews with the students
Novelty	Open schooling events	The sCYence fair was a lot of fun. The acting was fun. We had a lot of people to find out about this problem. I was going around asking people if they would like to come and I think there was only one person that knew what AMR was, everyone else said they didn't, so it was nice to get them to know what it is. (8RSR)
	Interaction with experts	I loved the trip to the biobank. It wasn't like every trip to the museum, boring. We had activities, it was interesting. We experienced some things we can't forget. (8PGMi) It was the first time that I interacted with scientists in their workplace. Apart from AMR, I learn about another study they do with public testing. Some people there were collecting urine samples to test for different things, there was a lab manager and some other people who worked to analyse different samples with microscopes. (8PKI)
	Scientific practices	I liked it because it felt more like we were really doing something instead of just learning. For example, we did research on the groups that we were representing and worked on real case studies, we had to look on proper medical websites. It felt like were not just kids in the class. It felt like we were actually doing a job. (8PKI)

The analysis of the qualitative data provides insights that support the results from the statistical analysis indicating the statistically significant difference with respect to question-item 2 'Doing science is one of my favourite activities'. It was evident that the open schooling events (i.e., science fair, awareness-raising campaign), the interaction with experts in an authentic environment as well as the engagement in scientific practices using authentic media and research equipment were perceived as an innovative feature in the school science. What becomes of interest here is that students' interactions with the experts provided the opportunity to enhance their understandings of science careers, how scientists work, what it takes to become a scientist and the skills required. This finding is consistent with the results of the statistical analysis indicating that students reported feeling 'very well informed' about the necessary steps for pursuing a

science-related career and the available opportunities in such fields.

Additional data from the interviews with the teachers support that these activities were interesting and useful to understand the science can offer for all the students regardless of their academic performance. This view is exemplified in the excerpt that follows from the interview with one of teachers:

> Certainly, the brightest students got a lot out of this. They put a lot in, and they put a lot out of it. The less able also enjoyed it and they were able to find a path through the project and the exercises and that it is more self-motivative for them. It's been quite a revelation how excited, interested and engaged the students have been in this, even students I wasn't expecting to be engaged. They obviously engaged in different degrees, but I think across the board the level of interest in the subject contents and the materials have been really very high. I would think that they got more engaged in science, more understanding of what science can offer and how broad science is, and they are more motivated in learning the subject. [..] I really liked the fact that we tried to introduce different experts to them, but I think because there were so many students, we couldn't really introduce the experts to them on a more personal level. I think it needs to be a little bit more concise, but they are way more aware about what scientists do than they were aware before because they got to see how they work in the lab, that was good for them. (Teacher A)

The words above illustrate that the students received career-related information, but more thorough planning needs to be considered for the main study by promoting a more personal (one on one) interaction with the experts.

Furthermore, responses placed under the category of knowledge (16) indicate that the students were more interested in sharing knowledge about AMR (9) than in learning new information with respect to the topic and the relevant careers (7). Finally, the less common category was social relatedness with students reporting being interested in group work (3) and in social interaction (2). Excerpts from the interviews with the students are some representative examples:

Table 5. Excerpts from interviews with the students indicate knowledge and social relatedness as features that enhance interest in science.

Category	Code	Excerpts from the interviews with the students
Knowledge	Learning/ knowledge gained	I liked learning lots of different things about antibiotics and resistance. The whole process was really fun and interesting. I loved every part of it. (8RSR)
	Sharing knowledge	I really felt that we were starting to show people what the problem was, and it was interesting to see that they did not really know anything and how they were interested in this. (8PKI)
	Raising awareness about AMR	Every single person I asked whether they knew about AMR, none knew, and it was good teaching them about this problem and ways to stop it. (8RSR)
Social relatedness	Working in groups	I liked all the projects we had to do because it was fun teaming up with other people. (8RSR)
	Social interaction	I liked the sCYence fair most because we displayed our posters and spread awareness about the issue while having fun and being able to do something interactive with the people. (8UDT)

Indicatively, the students were more interested in the novel activities than in learning and sharing knowledge. This implies that the students were engaged in scientific practices they were not familiar with in school science (e.g., argumentation and critical thinking tasks) and probably they were not feeling confident enough to reflect on the new knowledge and share it with other people.

Conclusions

The findings of this study suggest that open-schooling educational initiatives, as compared to traditional school science, contribute to students' perception of real science as more enjoyable, interesting, relevant, and informative with respect to careers, especially when they promote novelty, knowledge, and social connections. This pilot study showcased specific features that enhanced students' interest in science and career awareness, aligning with prior research (Drymiotou et al., 2021b) as follows: (a) organising open schooling events in the wider community; (b) interacting with experts in an authentic environment; (c) engaging in scientific practices; (d) promoting and sharing knowledge in general and specifically with respect to societal challenges; (e) promoting group work and social interaction.

Nevertheless, it is important to highlight that this pilot study has not investigated the role of the teacher in this intervention as well as the preparedness of the experts to act as science communicators. Hence, we recommend exploring these

two factors in the main study as they might influence the findings.

In conclusion, these findings hold important implications as they offer insights to: (a) inform the design of open-schooling teaching and learning activities; (b) further develop the open-schooling TLS framework, and (c) provide practical recommendations for curriculum design and classroom practices to enrich the curricula and pedagogical approaches in school science toward enhancing students' interest in science and awareness about science careers. One can argue that these activities opened students a window to real science providing connections between theory and practice.

Acknowledgement

This study has received support from the European Union through the Horizon2020 Research and Innovation program, project MULTIPLIERS, Grant agreement No 101006255.

References

Constantinou, C. P. & Papadouris, N. (2012) Teaching and Learning about Energy in Middle School: An Argument for an Epistemic Approach. *Studies in Science Education*, 48(2), 161-186.

Drymiotou, I. & Constantinou, C. P. (2023). Antimicrobial Resistance Toolkit. MULTIPIERS project. Nicosia. Accessed in https://multipliers-project.org/resources/.

Drymiotou, I., Constantinou, C. P., & Avraamidou, L. (2021a). Enhancing students' interest in science and understandings of STEM careers: the role of career-based scenarios. *International Journal of Science Education*, 43(5), 717-736.

Drymiotou, I., Constantinou, C. P., & Avraamidou, L. (2021b). Career-based scenarios as a mechanism, for fostering students' interest in science and understandings of STEM careers. *International Journal of Designs for Learning*, 12(3), 118–128.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences*, *2*(2), 141–178.

Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*. London: Sage Publications.

European Centre for Disease Prevention and Control [ECDC] (2022). *Antimicrobial Resistance in the EU/EEA - A One Health response*. Solna, Sweden: ECDC.

Kang, J., Keinonen, T., Simon, S., Rannikmäe, M., Soobard, R., & Direito, I. (2019). Scenario evaluation with relevance and interest (SERI): Development and validation of a scenario measurement tool for context-based learning. *International Journal of Science and Mathematics Education*, *17*(7), 1317–1338. https://doi.org/10.1007/s10763-018-9930-y

Merriam & Tisdell. (2015). *Qualitative Research: A guide to design and implementation.*

John Wiley & Sons.

Organisation for Economic Co-operation and Development (2021), *OECD Skills Outlook 2021: Learning for Life*, OECD Publishing, Paris. doi. 10.1787/0ae365b4-en

Papadouris, N. & Constantinou, C. P. (2017) Integrating the epistemic and ontological aspects of content knowledge in science teaching and learning. *International Journal of Science Education*, 39(6), 663-682. DOI: 10.1080/09500693.2017.1299950.

Papadouris, N. & Constantinou, C. P. (2016) Investigating middle school students' ability to develop energy as a framework for analyzing simple physical phenomena. *Journal of Research in Science Teaching*, 53(1), 119–145. DOI: 10.1002/tea.21248

UNESCO (2020). *Education for Sustainable Development - A Roadmap*, ESD for 2030; UNESCO Publishing: Paris, France.

"If you Really, Really, Want to Know it You Have to Touch it" – Researching the Role of Touch in Children's Science Education

Laura Colucci-Gray, Jonathan Hancock, Andrew Manches

Moray House School of Education and Sport, University of Edinburgh, UK

Whilst the sense of touch, or haptics, is commonly recognised as a crucial means of communication, social interaction, and sense-making, vital for exploring the world around us, its role in science education has been less clear, and less explored compared to written-visual and verbal-auditory modalities of learning. This likely reflects the dominance of cognitive paradigms in science education, which prioritise visual representations and conceptual abstractions, and largely underplaying the embodied nature of cognition grounded in movement and sensory experiences. Addressing this gap, this paper draws on a study involving 160 children aged 6-11 across six primary schools, based in Scotland, UK. Fieldnotes, video and audio recordings, and photographic data were collected during 47 observations of children interacting indoors with a series of commonly found natural artefacts, and over the course of eight outdoor workshops in the school grounds. Analysis revealed the complex ways in which younger learners experience, explore, and place value in touch, their awareness and readiness to touch in science learning, and the ways in which haptic experiences supported learning through action, verbal communication, and gesture. Findings disclose the less noticeable, affective and relational dimensions of scientific inquiry mediated by materiality and touch, and contribute to widening awareness of the value of sensorial inquiry in science education.

Keywords: Touch haptics, embodied learning, scientific inquiry.

Introduction

The sense of touch, or haptics, is widely considered vital to the human experience, forming an integral part of human multimodal system for understanding, communicating, and interacting in and with the world (Minogue & Jones, 2006). Contributions from the neurosciences associate tactile experiences with early language development, mediating the transfer of direct, sensory experience into abstract concepts and memory. For example, the feeling of 'hot' and 'cold', understood first as a tactile sensation through the skin, is essential for pattern recognition of how heat is transferred across surfaces, but it also carries an affective component, which may be used metaphorically to conceptualise other

situations, as in the common expression 'in the heat of things' or to refer to a relationship that may be 'going cold'. Arguably, early tactile experiences form a strong predictor of children's conceptual development and ability to make sense of phenomena later in life (Nathan, 2022).

Yet, despite emerging multimodal research on the role of touch in social and digital communication (Jewitt et al., 2022), there remains limited empirical evidence of the role of touch as a mode of science learning through which children explore, experience, understand, and value but also become (Kiverstein and Miller, 2015) in their environment. Accounting for this gap is both an empirical and a theoretical endeavour calling for an integrative and iterative approach to understanding the role of touch in cognition, by drawing on a wider range of cognitive theories addressing the body in learning. This study provides the initial findings of an investigation involving 160 participants aged 6-11 across six primary schools. Taking the lead from Kiverstein and Miller (2015), we look at touch from a wider perspective on embodied cognition, arguing for the inseparability of cognitive and emotional processing in the brain. In this view, children's engagement with tactile explorations of the world around them may be more fully understood and deployed educationally to support 'states of action readiness' (p. 1), involving the whole living body of the organism, and elicited by possibilities for action in the environment that matter to the organism. Such recognition becomes important when considering research which demonstrates that scientific literacy and expertise alone does not necessarily foster critical thinking or action (Sharon & Baram-Tsabari, 2020). At the same time, findings can further inform methodological and pedagogical approaches designed to augment the role of touch in learning and teaching, including haptic technologies for use in education.

Touch in Science Education: Moving from cognitive to ecological views of learning.

Studies on practical work in science and science education have typically framed the role of touch as aiding the manipulation of physical objects, to illustrate volumes, shapes, and function, aiding the transition from everyday experiences to scientific ways of thinking (Jones et al., 2006). For example, studies have demonstrated that sensory and touch interaction in the early years can provide a common experiential ground for children to learn and inquire together by sharing individual observations (Samuelsson, 2022). More recently, drawing largely from literature on informal learning and museum studies, the field of Objects Based Learning has focused on mapping the 'exploratory procedures' of touch to investigate how tactile interaction can support object identification and their classification (Lederman & Klatzky, 1987; Chatterjee et al., 2015), but also, how the physicality of the object can trigger and evoke emotional, physical and sensory responses with transformative value of their own (Dudley, 2012;

Novak & Schwan, 2021).

Emotional and affective encounters have been shown to be important aspects of science education, crucial in driving interest and engagement from students (Sinatra et al., 2014). This has been accompanied by increasing recognition of the 'perceptual power and the wealth of sensory information' afforded by the sense of touch (Jewitt et al., 2022). The latter is the central tenet of more recent contributions drawing on the expanding interdisciplinary field of embodied cognition, stretching from the early work of Merleau-Ponty (1962/2002) and Gibson's (1979) ecological theory of perception to more recent applications in neuroscience and neurolinguistics (Gallagher, 2014), as well as science education (Tamer et al., 2015). From an embodied perspective, cognitive processes are intrinsically linked to body-based interactions, and as such they are influenced by recursive sensory feedbacks between the organisms' body and their environment. In the categorisation provided by Wilson (2002), views of embodied cognition range from a more restricted focus on the environment acting in the form of an extended mind, aiding sense-making at the conceptual and abstract level, to more radical, ecological and less understood approaches, seeing organisms and environment as tightly coupled. At the one end of the spectrum, research in science education has evidenced the significance of physical artefacts – such as a step or a hill – acting as embodied metaphors which can be observed in language and gesture. For example, the arm stretching upwards may indicate an energy differential which is then reflected in language through images and metaphors (i.e., 'energy levels'; 'energy ladder') (Amin & Jepsson, 2015). From this perspective, embodied learning in science education establishes a close link between the subject and their cultural, social, and environmental context, and may support the teacher with greater understanding of the role of the physical environment as a mediator and predictor of children's engagement with science (Williams, 2012).

However, at the other end of the spectrum, post humanist and new materialist accounts emerging from quantum physics and epistemology of complexity in biology point towards an idea of the environment as part of human cognition itself, with an ongoing flow of materials, energy and information between mind and world being so dense and continuous that the mind alone is no longer a sufficient unit of analysis (Colombetti, 2014). This second view of embodied cognition is more firmly and clearly focused on the sensing body as an ecological nexus of mental and physical, porous but also bounded (Braidotti, 2013). This view decenters the human by shifting attention from questions about the who that makes and acts, to questions of how phenomena emerge from the embodied intra-actions of humans and more than human others (Colucci-Gray, 2023). Emerging literature on science education in the outdoors, for example, has highlighted modes of science learning that are cognitive, sensorial, and experiential, with

material encounters and affective responses prompting children's curiosity, engagement, and inquiry (Gray et al., 2021). Yet there remains a question as to how such processes can be integrated within a science education that accounts for synergistic and dynamic transformations of both organisms and their environments. There is therefore ground for furthering the scope of empirical investigations of touch as the gateway to children's first-person inquiries, with the capacity to provoke attentional states of knowing and being in the world (Haraway, 2016).

Theoretical Framework

Drawing together contributions from embodied cognition and post-human accounts (Braidotti 2013), this study considers human cognition as fundamentally enmeshed with objects, materials, and the environment. This notion decentres the human to foreground affective, sensorial, and experiential entanglements (Noë, 2023) which emerge as practices of affective knowing and care. From this position, we focus on touch as a process of cognitive unfolding, that is, the action with the potential to expose the child to something other than themselves while at the same time, establish continuity with past events and affective states embedded in memory. Arguably, from an embodied and enactivist perspective, these moments of intra-actions are at the root of growth and change, bringing the organism to a point where it is possible to acquire new experiences within the confines of what is known, while also disentangle oneself from what is known, in order to be able to perceive and integrate something new (Stendera, 2022). To undertake this inquiry, we adopt an integrative view of embodied cognition which, on the one hand, focuses on specific use of metaphors and gestures as consolidated, cognitive states and, on the other, is open to the emergence of new cognitive experiences, which are mediated by a renovated attentionality of the sensing body (van der Kamp et al., 2019). In practical terms, the analysis will combine audio and visual data to capture cognitive processes in action, focusing on touch as the interface between known and unknown, familiar and unfamiliar, abstract and concrete and vice versa.

Methodology and Methods

Design

This study is part of a larger, UKRI funded project (SENSE: Sensory Explorations of Nature in School Environments, EP/V042351/1) exploring the role and potential of natural and digital touch in education; the ways in which it may provoke different scientific questions and inquiries and help children to connect with the natural world.

Here we draw on a range of data sources from the project to present emerging

qualitative findings of the impact of tactile and sensory explorations on children's early science learning. Drawing on theories of embodied learning (Nathan, 2022) and post-humanism (Braidotti, 2013), three key principles guided the design and focus of the study:

1. Touch as part of the learning sensorium is an extension of the body-mind, hence, touching *is* knowing, and learning occurs at the point of difference between what may be expected and what is felt;
2. While visual perception is strongly linked to codified language, touch experience can help draw attention to the affective, experiential and contextual dimensions of knowing;
3. Touch experience mediates the imaginative and the metaphorical, expressed through physical movement and gesture (Brown, 2003).

Capturing and analysing the role of touch in children and young people's learning presents challenges due to age, prior experiences, and responsiveness to context (Ekström & Cekaite, 2020). Moreover, verbalising tactile experiences is a complex exercise, even for adults (Obrist et al., 2013). A semi-naturalistic approach was thus chosen for this study, combining observations indoors and outdoors across age groups and focussing on two central questions:

- **RQ1** In what ways are children ready to adopt and value touch experiences in their learning?
- **RQ2** In what ways are touch experiences mediating affective, material, and sensorial connections in science inquiry?

Stage 1 involved a sub-set of 100 children presented with a selection of commonly found natural artefacts across 47 observations. Sampling of schools was conducted across a broad socioeconomic range, as measured by the Scottish Index of Multiple Deprivation, to account for variety of contexts, experiences, and accessibility of science learning in the outdoors. A range of physical objects were selected as they were deemed to be familiar, or common: seashells, stones, a stick, tree bark, leaves, a bird feather, a pinecone, a fossil, a shark tooth, and a flower. Common natural objects were chosen to investigate the extent to which touching and manipulating revealed 'new' or 'unknown' features to the children, that would not have been discernible by using vision alone. It was also acknowledged that while the objects may appear familiar to many children, they may be less familiar to others: given the socioeconomic context of some schools, the children may have had less experiences in the outdoors and in nature, which was evident in observations when some children commented on never having been to the beach or touched certain objects, for example the pinecone. Natural objects were also chosen for the variety of texture and the tactile sensation and touch exploration they could offer (Lederman & Klatzky, 1987).

Table 1. Participants.

	Indoor Observations	No. of Children	Outdoor Workshops	No. of Children	Age
School 1	7	19	1	24	8-9
School 2	6	12	2	22	6-7
School 3	8	16	2	47	6-7
School 4	8	16	1	26	7-8
School 5	8	17	1	17	8-11
School 6	10	20	1	24	8-9
Total	47	100	8	160	

Observations took place throughout the school day, lasting for 15 to 20 minutes on average, to allow time for introductions, building rapport and communicating trust, and for the children to partake in the activities and ask questions about the research. The aim was to establish a natural but structured environment through which variation of touch interaction, perspectives, and experiences were allowed to emerge. Semi-structured interviews layered in prompts at pre-determined intervals, with a script that gradually scaffolded children towards tactile interaction, allowing for activities to be partly child-led and for researchers to observe how naturally children used touch for learning. Responses were captured through real-time observation, and video and audio recordings where consent was given. Real-time observation was necessarily less detailed but enabled capturing information on the frequency and typologies of touch, manipulation as well as gesture and speech across the different groups. Video and audio recording supported a more in-depth analysis, as well as triangulation and validation of real-time observational data.

Stage 2 involved 160 children taking part in a 2-hour workshop at their school, led by an external outdoor learning specialist. In the workshops, children explored their school grounds, engaging in sensory and tactile experiences, and interpreting and communicating these experiences through creative and artistic activities including clay modelling, printing, drawing, painting, and displaying. Video clips, images, and observation notes focussed on salient episodes of tactile interactions and artistic interpretation, and the beginnings of scientific inquiry (Brown, 2003).

Analysis

In the analysis, touch, and the communication and articulation of touch, were understood as multimodal (Jewitt et al., 2022), recognising the role of language, gesture, and tactile interaction. First, qualitative thematic multimodal analysis of the data was employed in three stages:

- Audio recordings transcribed and emerging themes coded.

- Video recordings and observation notes matched for accuracy.
- Notes and coding from transcripts were layered on observation schedules.

From this initial analysis five key dimensions were identified, representing different modes of touch communication and experience and tactile inquiry: (i) Non-verbal touch interactions; (ii) Non-tactile interactions; (iii) Affective encounters; (iv) Tactile gestures; (v) Sensorial and scientific inquiry. Drawing on these dimensions and applying a lens of post-humanism and socio-materiality in science education across the data, two key and interrelated themes emerged: 1) Inquiry led by materiality, and 2) Inquiry led by affect.

Ethics

Consent was obtained in advance from parents and carers in collaboration with the schools. Children were asked to give verbal consent before the observation and recording was started. They were assured that they would be allowed to participate in the activities even if they decided they did not want to be recorded.

Findings

Findings revealed a nuanced picture of how touch played a role in children's exploration and communication of their ideas, as well as more emotional aspects. These are reported here to evidence the varied relationship between cognition and touch modalities.

Inquiry led by materiality

In the indoor observations, children's propensity to touch objects varied significantly. Episodes of talk without touching often revealed the influence of prior learning as well as earlier touch experiences. For example, these were evoked in terms of perceived similarities as to what an object might feel like by the way it looked: "I would say if I described this one it's kinda, not fluffy, but, I don't know" [sitting back from table, looking at objects] (Child, age 9, School 1, Observation 1). Children also expressed the difficulty of describing something without touching.

Across the study, when children were touching and manipulating, they showcased a greater level of description and interest in the natural objects they were interacting with. Metaphors, comparisons, and analogies were used to describe patterns of variation and commonality across different objects, and notably, such linguistic expressions were accompanied by gestural motions of the fingers literally 'sampling and savouring' the material at hand. In the following exchange from School 2, Observation 2, the shared material encounter with an

object, a shark tooth, opened up a dialogic space for the children to explore and investigate through touch:

> **Child 2:** it's quite dusty [*gesture, rubs thumb and index finger together]* … This bit feels smooth *[rubbing thumb on object]*, and this bit feels a bit like...
>
> **Child 1:** Yeah *[takes object from Child 2]*. And this bit feels like *[pause]*… Rock. Maybe, well, I think, *[showing to Child 2, speaking in lower tone]* look because that bit's like smoother, it looks like it'll come up here *[running thumb and finger along object]*.
>
> **Child 2**: Yeah, it's quite smooth *[rubbing object with fingers and thumbs]*.
>
> **Child 1:** Maybe it's a bit like the layer of the tooth has just broken off or something *[gestures hand pulling away, as if breaking something off]*.

Gestures were used often by children to describe or recall a tactile experience or sensation; gesturing became a mode through which they could communicate touch experience and articulate hypotheses or questions they had about the natural objects and their material features: "Maybe it's a bit like…." (Child 1). Significantly, the appearance of the pauses also point to a slowing down in the encounter (Haraway, 2016) and discovery of texture and materiality, through which tactile language and gesture is developed and the foundations of inquiry are established.

In some cases, touch involved a focused, intensive manipulation, drawing the object to themselves, with eye gaze firmly on the object, sometimes held close to the eye or face (Fig. 1a). This was also drawn out in the outdoor workshops, and across the age groups. Close, sometimes silent, concentrated eye gaze and physical manipulation of the objects (Fig. 1b and 1c) emerged as a feature of children's inquiry through touch, followed by a move towards voicing discoveries. Some children were vocal about the questions that the tactile sensations of the objects were prompting: "It's all mossy and interesting [turning back around, rubbing finger on surface] and like, what tree is it from? How old is it? [gesture, rubbing thumb on finger]… there's a lot of questions behind these things if you really think" (Child, age 8, School 4, Observation 1).

Figures 1a, 1b, and 1c. Silent manipulation; drawing object close to face; tactile gesture.



In outdoor workshops, many children first described the objects they found in the school grounds using visual vocabulary, for example focusing on the colours of leaves. Once engaging with the tactile and material features, and through artistic or creative activities, there was growing recognition and awareness of the variety and diversity of natural objects. At School 5, one child (aged 10) took bark rubbings of all the trees on the school grounds (Fig. 2a), identifying differences in texture and posing questions about the bark. At the same school, leaf prints (Fig. 2b) allowed children to notice differences in texture and tactile sensations of certain leaves and an understanding that they are not homogeneous, "it's soft, but it's also hard here… and kinda prickly" (Child, aged 9, School 5). At School 3, one boy took samples of different types of bark (Fig. 2c), displaying them on card according to texture: "this is kinda smooth… this bark is squishy… this one is kinda hairy" (Child, aged 8, School 3, Outdoor Workshop). The boy presented the display to the rest of the class, started calling himself 'bark boy', and asked questions of the teacher about the nature and texture of the barks. Here an interest and excitement sparked by an encounter with the materiality of natural objects in the outdoors provided a platform for child-led scientific inquiry.

Figures 2a, 2b, and 2c. Tree bark rubbings; leaf print; bark display.



Inquiry led by affect

Integrated in the development of inquiry led by materiality, there appeared to be a notable affective dimension to children's tactile exploration. In the indoor

observations, most of the children at some point engaged in non-verbal touch, manipulating an object without directly looking at it; often this was a natural object, such as the feather, but at times it was an item of clothing, for example rubbing the sleeve of their jumper. Here, the child's eye gaze was directed elsewhere, sometimes at the interviewer asking the question, or at the other objects on the table. The tactile sensation in this case appeared to offer comfort to the child, in the same way that touch is often used as a means of emotional regulation in children (Cekaite & Holm Kvist, 2017). Touch here appears to have played a significant role in the child attending to their own feelings and emotions, allowing space and time for them to engage in the activities provided.

When interacting with the objects, children commented on the 'calming' nature of touch, or explained enjoyment of the different textures and tactile sensations, as in the following quote:

> I like twisting it and feeling like the different parts of the feather… there's a wee bit of a feather kind of bit here, and then like a wee stem… kind a like a tree or flower how it has the stem. And then, the delicate part… I wonder if the feather feels the same as the flower? Feel the flower. *[handing to other child]* (**Child, age 9, School 1, Observation 2**).

In this instance, touch affords noticing deeply (van der Kamp et al., 2019). The inquiry is activated first by an affective encounter, the child enjoying the tactile sensation of the feather, during which variety of the object is acknowledged along with its 'delicate' features; it also prompts the comparison of the texture of the feather with that of the flower, establishing a 'shared experiential ground' between the two children (Samuelsson, 2022). At other times, children recounting memories prompted by tactile sensation appeared to have emotional connotations; for example, touching the shell recalling a family holiday, or the tree bark texture surfacing memories of a treehouse in a previous home (**Child, aged 8, School 4, Observation 3**).

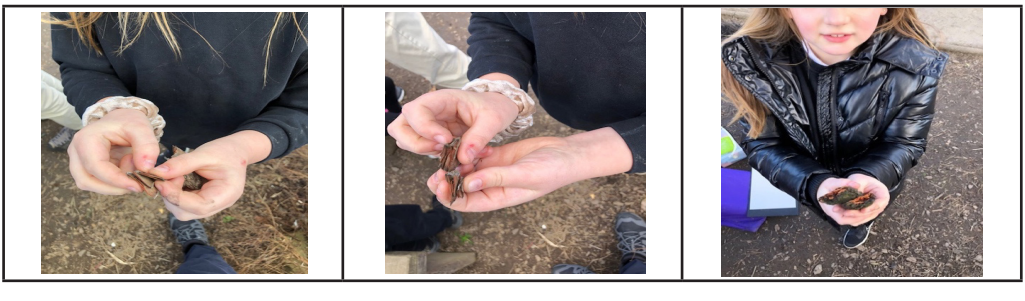
Not all affective encounters in the indoor observations were positive; some children spoke about disliking certain textures or experiences, or finding them 'uncomfortable' or 'scary'. However, tactile sensations deemed unpleasant could still prompt investigation and inquiry. For example, a girl in School 6, Observation 2 ran her fingers over a leaf, describing it as "eww… wrinkly", with a facial expression indicating the experience was unpleasant. The girl then immediately picked up the other leaf on the table, eager to compare: "it's way smoother [*rubbing thumb and fingers together in tactile gesture*] … that one's so smooth" (**Child, aged 8, School 6, Observation 2**).

In the outdoor workshops, the children's interest in and description of artefacts was often emotional, both with positive associations ("it's so soft") and anticipating potentially negative ones, for example touching a nettle ("I'll get

stung"), holly branch ("it's spikey"), or fungi ("it looks slimy") (School 3, Outdoor Workshop). Some children picked up on the 'comforting' nature of some textures. For instance, in School 5, an 11-year-old girl described the leaf she found as a "soft jumper leaf": "I picked a really smooth leaf, it doesn't look that smooth from far away… If you shut your eyes you wouldn't know it's a leaf."

At School 4, situated in an area of high deprivation, and with a largely concrete school ground with little nature or biodiversity, several children did not recognise or commented that they had never touched some of the 'common' natural objects. In the outdoor workshop, the children displayed surprise and curiosity at the variety of textures. Some of the children were protective of the nature objects they had found, including two 8-year-old girls who had collected bark: "I need to keep mine safe, I don't want to lose them" (Child, aged 8, School 4, Outdoor Workshop). The girls proceeded to break up and snap the bark into pieces, and were surprised and excited to find the layers inside: "Oh wow! Look at the inside of this, it's different inside, it's layers… it feels different [running finger along] … I love breaking this" (Figs 3a, 3b, 3c). The educator explained that the touch sensation of the bark is due to water getting inside and talked about how the texture of bark changes when exposed to water. The affective encounter of breaking the bark and touching the inside therefore afforded the educator and learners opportunities for science learning and inquiry; tactile exploration here opens dialogic spaces (Braidotti, 2013), between human and human, and between human and more than, and through this space allows learners to 'meet the world as it is' while developing scientific knowledge, curiosity, and wonder.

Figures 3a, 3b, and 3c. Breaking up the bark.



Discussion and Implications

Echoing prior literature (Novak & Schwan, 2021; Minogue & Jones, 2006), this study demonstrates the continued dominance of children's visual modality in their mode of knowing, often superseding tactile modalities. Yet, the opportunity to analyse tactile interaction in more detail revealed a much more complex and significant role of touch in children's engagement, exploration, and communication of science ideas through both words and gestures. Indeed, an emerging implication of this work is the need to draw children's and educators' attention to the significance of touch as a legitimate and direct modality through which children orient themselves in the world.

In the indoor observations and outdoor workshops, the materiality of the natural objects and the affective encounters mediated by touch prompted curiosity and inquiry in ways that using vision alone did not. Materials that may at first have appeared 'familiar' to some children became 'unfamiliar', evoking a range of responses and questions which allowed educators to build on and connect to scientific concepts and ideas. Creating a dialogic space for these experiential, sensorial, and affective encounters through tactile interaction allowed children to articulate nuanced touch experience through multimodal means, something that has proved challenging to capture in previous research (Jewitt et al., 2022). It is also important to reflect on the role that artistic and creative activities may allow children to interpret and communicate their touch experiences in varied ways and forms, enhancing tactile science inquiry.

Importantly, the use of gestures to describe abstract ideas may indicate that children have internalised previous tactile sensations, which they draw upon to ground their experiences, engage and share in meaning-making, and form conceptual understandings (Nathan, 2022). Further investigation and categorisation of these gestures may provide insights into how these are employed for learning, especially for children new to the school setting, or those who may not be readily exposed to learning outdoors, as observed in some of the schools of greater economic deprivation. Findings point to new avenues through which educational researchers can understand the ways in which touch speaks of the experiences of children, and the ways in which they learn science and about their presence in the world (Braidotti, 2013) through touch, which will likely vary across communities, contexts and domains of learning.

Acknowledgement

We would like to place on record our sincere thanks to the children, teachers, and schools who took part in the study, and the outdoor learning specialist who organised and conducted the outdoor workshop.

References

Amin, T., & Jepsson, F. (2015). Conceptual metaphor and embodied cognition in scinece learning: Introduction to special issue. *International Journal of Science Education*, *37*(5-6), 745-758.

Braidotti, R. (2013). *The Posthuman*. Polity Press.

Brown, T. L. (2003). *Making Truth. Metaphors in Science*. Illinois Press.

Cekaite, A., & Holm Kvist, M. (2017). The comforting touch: tactile intimacy and talk in managing children's distress. *Research on Language and Social Interaction, 50*(2), 109–127.

Chatterjee, H. J., Hannan, L., & Thomson, L. (2015). An introduction to object-based learning and multisensory engagement. In H. J. Chatterjee & L. Hannan (Eds.), *Engaging the senses: Object-based learning in higher education* (pp. 1–18). New York: Routledge.

Colombetti, G. (2014). *The Feeling Body: Affective Science Meets the Enactive Mind.* Cambridge, MA: MIT Press.

Colucci-Gray, L. (2023). Gesturing in plain sight: dialogical enactments of sustainable futures as being and doing in the world. *Cult Stud of Sci Educ*. https://doi.org/10.1007/s11422-023-10189-w

Dudley, S. (2012). *Museum objects: Experiencing the properties of things*. London: Routledge.

Ekström, A., & Cekaite, A. (2020). Children's touch in a Swedish preschool: touch cultures in peer group interaction. *International Journal of Early Years Education*.

https://doi.org/10.1080/09669760.2020.1857709

Gallagher, S. (2014). Phenomenology and embodied cognition. *The Routledge handbook of embodied cognition, 9-18*.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Houghton, Mifflin and Company.

Gray, D., Colucci-Gray, L., & Robertson, L. (2021). Cultivating primary creativities in STEAM gardens. In P. Burnard, & M. Loughrey (Eds.), *Sculpting New Creativities in Primary Education* (1st ed., pp. 146-161). Routledge. https://doi.org/10.4324/9781003129714

Haraway D. (2016). *Staying with the trouble: Making kin in the chthulucene (Experimental Futures)*. Duke University Press. https://doi.org/10.2307/j.ctv11cw25q

Jewitt, C., Price, S., Atkinson, D., & Golmohammadi, L. (2022). Editorial. *Multimodality & Society*, *2*(3), 203–212. https://doi.org/10.1177/26349795221113038

Jones, M. G., Minogue, J., Tretter, T. R., Negishi, A., & Taylor, R. (2006). Haptic augmentation of science instruction: Does touch matter?. *Sci. Ed., 90*, 111-123.

Kiverstein, J., & Miller, M. (2015). The embodied brain: towards a radical embodied cognitive neuroscience. *Front. Hum. Neurosci*., *9*, 1-11.

Lederman, S. J., & Klatzky, R. L. (1987). Hand movements: a window into haptic

object recognition. *Cognitive Psychology*, *19*(3), 342–368. https://doi.org/10.1016/0010-0285(87)90008-9.

Merleau-Ponty M. (1962/2002). *Phenomenology of perception*. London, UK: Routledge.

Minogue, J., & Jones, M. G. (2006). Haptics in education: exploring an untapped sensory modality. *Rev. Educ. Res. 76*(3), 317–348.

Nathan, M. J. (2022). *Foundations of Embodied Learning: A paradigm for education*. New York: Routledge.

Noë, A. (2023). *The Entanglement: How art and philosophy make us what we are*. New Jersey: Princeton University Press.

Novak, M., & Schwan, S. (2021). Does Touching Real Objects Affect Learning? *Educ Psychol Rev, 33*, 637–665.

Obrist, M., Seah, S. A., & Subramanian, S. (2013). Talking about tactile experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. Association for Computing Machinery, New York, NY, USA, 1659–1668.

Samuelsson, R. (2022). Touch and translanguaging in a multilingual early childhood education setting. *Multimodality & Society*, *2*(3), 300–322. https://doi.org/10.1177/26349795221109355

Sharon, A. J., & Baram-Tsabari, A. (2020). Can science literacy help individuals identify misinformation in everyday life?. *Science Education*, *104*(5), 873-894.

Sinatra, G. M., Broughton, S. H., & Lombardi, D. (2014). Emotions in science education. *International handbook of emotions in education*, 415-436.

Stendera, M. (2022). Beyond Disintegration: Transhumanism and Enactivism. In E. Tumilty, E. & M. Battle-Fisher (Eds.), *Transhumanism: Entering an Era of Bodyhacking and Radical Human Modification. The International Library of Bioethics, vol 100*. Springer, Cham.

Tamer, A G., Jeppsson, F., & Haglund, J. (2015). Conceptual Metaphor and Embodied Cognition in Science Learning: Introduction to Special Issue. *International Journal of Science Education, 37*(5–6), 745–758.

Williams, R. F. (2012). Image schemas in clock-reading: Latent errors and emerging expertise. *Journal of the Learning Sciences*, *21*(2), 216–246.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625-636.

van der Kamp, J., Withagen, R., & Orth, D. (2019). On the Education About/of Radical Embodied Cognition. *Front. Psychol*., *10*, 1-9.

Examination of Effects of Authentic Scientific Inquiry Experiences on Students' Epistemological Beliefs

Ozlem Goreci, Feral Ogan-Bekiroglu

Marmara University, Turkiye

The development of sophisticated epistemological beliefs about science is an essential goal of science education and overall science literacy. The purpose of this study was to examine the effectiveness of authentic scientific inquiry on students' epistemological beliefs about science. Models-of-data theory framed this research. A pretest/posttest control group design was conducted to monitor the change in students' scientific epistemologies and to measure the effects of implementation of authentic scientific inquiry. The students' epistemological beliefs about science were measured with self-report questionnaire. Results showed that the students became more sophisticated in their beliefs about source and certainty of scientific knowledge and development of knowledge dimensions. However, there was not any significant change in the students' justification of scientific knowledge beliefs. This research suggests that more studies focusing on students' epistemic sophistication with the help of authentic scientific inquiry experiences would be done.

Keywords: Scientific epistemological beliefs, authentic scientific inquiry, physics.

Introduction

Epistemological belief or personal epistemology is a psychological construct that refers broadly to individual conceptions and theories of the nature of knowledge and knowing (Burr & Hofer, 2002). According to Hofer (2004), personal epistemology is a set of beliefs, is multidimensional, develops over time and in relation to education and experience, exists at both the domain-general and domain-specific level, and is activated in context. Hofer and Pintrich (1997) proposed four dimensions for epistemological beliefs and they clustered personal epistemology into two areas: the nature of knowledge (what one believes knowledge is) and the nature or process of knowing (how one comes to know). Within these, there appear to be two dimensions each. Under the nature of knowledge area, there are the dimensions of certainty of knowledge and simplicity of knowledge. Within the area of nature of knowing, there are source of knowledge and justification for knowing dimensions. Epistemological beliefs can influence both the kinds of new information that is picked up from the physical and sociocultural context and the way in which this information is

interpreted (Stathopoulou & Vosniadou, 2007).

Science as a socially situated practice in which scientists' values for what count as good questions, appropriate methods, and good answers are constructed and negotiated within particular scientific disciplines and communities (Sandoval & Morrison, 2003). Scientific epistemology refers to students' beliefs and views about how scientific knowledge is developed and justified, and involves a set of ideas and assumptions about the nature of science that students have (Wu & Wu, 2011). The development of sophisticated epistemological beliefs about science is an essential goal of science education and overall science literacy (Peffer & Ramezani, 2019).

How "inquiry" gets implemented in particular classrooms has direct consequences upon the epistemological ideas (Sandoval, 2005). Abd-El-Khalick and his colleagues (2004) argue that when students learn science in the inquiry context, they develop epistemological understandings about nature of science and scientific knowledge. However, little work has explicitly addressed the relation between students' scientific epistemologies and their inquiry (Sandoval & Morrison, 2003) and the results are controversial. A number of studies in students' epistemologies indicated that although in inquiry-based classrooms students engage in activities that share similarities with those done by practicing scientists, students still hold naïve epistemological views and their epistemological ideas seem difficult to change through inquiry practices (Wu & Wu, 2011). For example, Sandoval and Morrison (2003) explored high school students' ideas about theories and theory change. They revealed that students' epistemological ideas did not change as a result of their inquiry experiences, but they could distinguished ideas from experiments and saw the purpose of experimentation as testing ideas. Zhao et al. (2021) explored the effects of prediction-observation-explanation inquiry-based learning model on fifth graders' ccientific epistemological beliefs and showed significant increase in experimental group learners' beliefs in the source and certainty dimensions.

Authenticity as a crucial concept to motivate students for science, to help them to connect scientific concepts and transfer knowledge, and to achieve more positive attitudes towards science. the concept of authenticity is multidimensional (Schriebl, Müller & Robin, 2023). Science fairs, research apprenticeship programmes, argument-driven inquiry, and model-based inquiry are among the most authentic experiences for K-12 science learners (Burgin, 2020; Windschitl, Thompson & Braaten, 2008). Authentic experiences are 'coherent, meaningful and purposeful activities defined as ordinary practices of the culture' (Brown et al., 1989, p. 34). The most common way of defining authenticity in science is providing students the experience "of what scientists do, how science is done,

and what science is" (Rowland et al. 2016, p. 5). In the context of authentic science inquiry environment how a student engages with source (where do participants look for information, and why?), justification (how data are used to support claims?), certainty (how do the students discuss their results?), and development (how does the student's interpretation of the problem change in light of new information?) can be observed and analysed (Peffer & Ramezani, 2019).

The epistemology that underlies simple inquiry tasks is very different from the epistemology that guides authentic scientific reasoning (Chinn & Malhotra, 2002). Sandoval (2005) suggests that engaging students in authentic scientific inquiry is an excellent pedagogical approach for promoting students' content learning, inquiry skill development, and understanding of the nature of science and of scientific models. On the other hand, Chinn and Malhotra (2002) found that textbook activities have almost no epistemological authenticity, and research-based efforts focus on a fairly narrow bandwidth of epistemological features related to generating and interpreting data. Sandoval and Morrison (2003) and Sandoval and Reiser (2004) suggested an epistemic discourse or using epistemic tools to relate students' inquiry experiences directly to the practice of professional science to support their development of more sophisticated scientific epistemologies. There is not much attempt made to improve students' scientific epistemologies via authentic science experiences including epistemic features. Therefore, the purpose of this study was to investigate the effectiveness of authentic scientific inquiry on students' epistemological beliefs about science.

Theoretical Framework

Models-of-data theory developed by Chinn and Brewer (2001) framed this research. Models-of-data theory helps us understand the difference between authentic inquiry and simple classroom inquiry. Authentic scientific inquiry refers to the research that scientists actually carry out. There are epistemological differences between simple inquiry tasks and authentic scientific inquiry in terms of purpose of research, theory-data coordination, theory-leadenness of methods, responses to anomalous data, nature of reasoning, and social construction of knowledge (Chinn & Malhotra, 2002). According to Chinn and Brewer (2001), the models-of-data framework points to the importance of developing more complex tasks on which students can learn to evaluate evidence. Authentic inquiry in science typically involves data and theories integrated in models that have a high degree of complexity (Chinn & Brewer, 2001).

Methodology

A pretest/posttest control group design was conducted to monitor the change in students' epistemological beliefs about science and to measure the effect of implementation of authentic scientific inquiry. Participants of this study were 39 ninth grade students studying in an urban high school. The research was conducted during the physics course and lasted ten weeks including application of pre-and post-tests. The experimental group was randomly selected by drawing lots between two classrooms. The instructional strategy in the experimental group was based on authentic scientific inquiry while curriculum-based instruction was followed in the control group. The students took the physics course two hours a week and worked as groups when it was necessary. The same learning objectives were pursued in both groups. The students in both groups studied the same energy concepts with the same teacher. Ninth-grade physics textbook activities were implemented in the control group. However, these activities were modified to become more authentic by adding extra questions and epistemic inquiry procedures in the experimental group. Burgin (2020) describes the top level of authenticity as follows: A learner-developed investigation that is interesting/meaningful, relevant, and makes a widely shared contribution to others. Therefore, the students in the experimental group answered the following questions during the experiments (Bybee, 2002):

- What is known and what questions are raised by this knowledge?
- What investigative procedure will address particular questions?
- What quipment is necessary to carry out this procedure?
- What data relevant and should be collected?
- How will the data be analyzed?
- What does the data mean?
- What mathematical calculations, if any, are required and in which order should they occur?
- How is the work to be communicated to others?

The teacher encouraged the students in the experimental group to record explanations, use data to support their claims, and justify why selected data were good evidence during each experiment. In addition, student groups reviewed each other's explanations (Sandoval & Morrison, 2003). While the students in the experimental group generate research questions, select variables to investigate, and employ multiple forms of argument; research questions are provided to the students in the control group, they investigate one or two variables and employ simple contrastive reasoning. The experimental group started activities with scenarios or problems whereas the control group followed the directions. The

differences between the experimental and control groups in terms of dimensions of epistemology are presented in Table 1 (Chinn & Malhotra, 2002).

Table 1. Epistemology Dimensions in the Groups

Dimesion of Epistemology	Experimental Group	Control Group
Purpose of Research	Students aim to build and revise theoretical models.	Students aim to uncover a simple surface-level regularity.
Theory-Data Coordination	Students coordinate theoretical models with multiple sets of data.	Students coordinate one set of observable results with conlusions about those observable results.
Theory-Ladenness of Methods	Methods are partially theory laden.	Methods are not theory laden.
Responses to Anomalous Data	Students rationally and regularly discount anomalous data.	There is little scope for students to rationally discount data.
Nature of Reasoning	Students employ heuristic, nonalgorithmic reasoning. Students employ multiple acceptable argument forms. Reasoning is uncertain.	Students employ algorithmic reasoning to derive a conclusion from an experiment. Students employ simple contrastive arguments. Reasoning is certain.
Social Construction of Knowledge	Students build on previous research by many scientists. Institutional norms are established through expert review processes and exemplary models of research.	Students seldom build on any previous research. There are no institutional norm-setting processes.

Sandoval (2005) called students' beliefs or ideas for professional science as formal epistemology. Interviews and questionnaire are popular assessments to explore students' formal epistemologies. The participants' epistemological beliefs about science were measured with self-report Scientific Epistemological Beliefs (SEB) questionnaire developed by Conley, Pintrich, Vekiri and Harrison (2004). It was a 26-item instrument and the items were rated on a 5-point Likert scale and distributed under the following four dimensions: source, certainty, development and justification. The source and certainty dimensions were reverse coded. Higher scores indicated more sophisticated scientific epistemic beliefs. For each domain of SEB, individual items were averaged together to create a single score for each of the four domains. Total score on the scientific epistemological beliefs survey was found by calculating the average across all items.

According to Conley and his collegues (2004), a less sophisticated belief

about the source of science knowledge is that it is derived from an outside authority person, rather than by resulting from one's inquiry strategies. They argue that justification in science is how an individual uses data or evidence, particularly generated through experiments, to support their claims. Certainty, on the other hand, refers to the nature of knowledge as concrete versus tenuous. Someone with a less sophisticated understanding of certainty would claim that scientific knowledge is certain and unchanging and that it is possible to obtain a single "correct" answer (Conley et al. 2004). This belief also relates to the development domain in that a more sophisticated understanding would be that scientific information could change in light of new developments (Conley et al. 2004).

The participants' mean values are divided into three categories as beginning (M = 1.00 – 2.33), intermediate (M = 2.34 – 3.67), and sophisticated (M = 3.68 – 5.00) levels. Wu and Wu (2011)'s explanations for each level are taken into account. At the beginning level, students cannot differentiate between evidence, hypotheses, and theories. They focus on procedures and results, and have a strong belief in the certainty of scientific knowledge (i.e., scientific knowledge is unproblematic and unchanging) (Wu & Wu, 2011). At the intermediate level, students start making an initial differentiation between scientists' ideas, procedures, and results; they define experiments as tests of hypotheses, view theories as well-tested hypotheses, and believe that knowledge is uncertain. Additionally, students realize that scientists design experiments to test their initial ideas and that one purpose of science is to understand questions about why something happens (Wu & Wu, 2011). At the sophisticated level, students view a scientific theory as an explanatory framework and make a distinction between a theory and a hypothesis that can be tested within a theory. Students realize the theory-guided nature of scientific inquiry and view knowledge as uncertain (i.e., scientific knowledge is tentative and can be revised and changed as new understandings are learned) (Wu & Wu, 2011).

Ozkan (2008) adapted the EBS for seventh-grade students into Turkish and two factors (sources and certainty) from different areas merged into one factor. Internal consistency computed by the Cronbach's alpha was good with reliability coefficients of 0.75 for the pre-test and 0.79 for the post-test for this study. Shapiro-Wilk tests were performed for normality analysis. Effect sizes were calculated for the changes in the groups. Independent t-tests were performed between the groups to make comparison and paired t-tests were used to analyze the data within groups.

Results and Discussion

Results of independent group t-test analysis presented in Table 2 indicate that there was no significant differenve between the control and experimental groups'

epistemological beliefs about science levels before the instruction ($t_{(37)}$ = 1.20, p > 0.05). However, the mean score of the students in the experimental group was significantly higher (M = 3.66, SD = 0.38) than the mean score of the students in the control group (M = 3.34, SD = 0.50) after the instruction ($t_{(37)}$ = -2.16, p < 0.05). The effect size between the post-tests of the groups was calculated as (d = 0.69). Accordingly, the change between the post-tests of the groups was moderate and that 69% of this observed change was due to the intervention. That is, authentic scientific inquiry activities made positive changes in the students' epistemological beliefs about science.

Table 2. Results of Independent Group t-Test Analysis for Epistemological Beliefs about Science

Variables	N	M	SD	Min Score	Max Score	df	t	p
Control Group Pre-Test	*20*	3.15	0.32	2.62	3.69			
Experimental Group Pre-Test	*19*	3.02	0.33	2.38	3.77	37	1.20	0.237
Control Group Post-Test	20	3.34	0.50	2.38	4.12			
Experimental Group Post-Test	19	3.66	0.38	2.77	4.27	37	-2.17	0.036

Table 3. Results of Paired Group t-Test Analysis for Epistemological Beliefs about Science

Variables	N	M	SD	Min Score	Max Score	df	t	p
Control Group Pre- test	20	3.15	0.32	2.62	3.69			
Control Group Post- test	20	3.34	0.50	2.38	4.12	19	-1.58	0.12
Experimental Group Pre- test	19	3.02	0.33	2.38	3.77			
Experimental Group Pre- test	19	3.66	0.38	2.77	4.27	18	-5.22	0.00

Regarding the results of paired group t-test analysis given in Table 3, while the increase in the control group's mean scores was not significant ($t_{(19)}$ = -1.58, p > 0.05), the increase in the experimental group's mean scores was significant ($t_{(18)}$ = -5.22, p = 0.00). The studenst who performed textbook activities could not show any development in their epistemological beliefs about science because these activities have almost no epistemological authenticity, and research-based efforts focus on a fairly narrow bandwidth of epistemological features related to generating and interpreting data (Sandoval, 2005). The pretest and posttest effect size of the experimental group was calculated as (d = -1.2). The change in the experimental group is seen at a high level. Although the students in the experimental group could not raise their scientific epistemology from

intermediate level to sophisticated level, they showed improvement. More complex activities done in the experimental group might afford greater use of authentic scientific reasoning processes and thus encouraged a more scientific epistemology (Chinn & Malhotra, 2002). The results of this current study are in line with Schiefer et al. (2020)'s results because they indicated that the children assigned to the extracurricular STEM enrichment programme developed more sophisticated epistemic beliefs and a higher level of epistemic curiosity than the children assigned to the control condition. Similarly, Shi, Ma and Wang (2020) found that students in a traditional cookbook guided laboratory showed significant negative shifts on personal epistemologies, and in contrast, students' epistemologies of experimental physics in inquiry-based laboratory were significantly improved.

Table 4. Results of Paired Group t-Test Analysis for the Experimental Group's Epistemological Beliefs about Science Sub-Dimension

Variables	N	M	SD	Min Score	Max Score	t	df	p
Development Pre-test	19	3.41	0.58	2.00	4.50			
Development Post-test		3.80	0.66	2.33	5.00	-1.83	19	0.08
Source and Certainty Pre-test		2.51	0,45	1,55	3,18			
Source and Certainty Post-test	19	358	0.44	2.36	4.18	-7.23	19	0.00
Justification Pre-test	19	3.39	0.61	2.33	4.89	-1.66	19	0.05
Justification Post-test		3.65	0.61	2.22	4.56			

The results of paired samples t-tests revealed that there was a significant increase in the experimental group's beliefs in the source and certainty dimension from pre-test (M = 2.51, SD = 0.45) to post-test ((M = 3.58, SD = 0.44), t(19) = -7.23, p = 0.00) (see Table 4). Effect size for this increase was medium (d = 0.66). The students became more sophisticated in their beliefs about source and certainty of knowledge dimensions after having experiences with authentic science inquiry. That is, there was a shift in the students' perception of self as an active maker of meaning and they began to see knowledge as an intuitive reaction, personally experienced. They did not perceive authority as the only source of knowledge and mostly held their own opinions as equally valid. There was a growing realization among the students that one cannot know with certainty and knowledge can be changed over time. However, there was not any significant change in the students' justification of knowledge beliefs. This dimension is concerned with the ways in which students use evidence and evaluate claims. Authentic scientific inquiry did not make any improvements in the students' concerns about the role of experiments and the use of data to support arguments. In addition, the progress in the development of knowledge dimension was not significant either. The students could not completely recognize that science is an evolving subject and that ideas and theories can change on the basis of new

data and evidence. The teacher can provide opportunities for students to notice when claims are unsupported and when causal mechanisms are not clearly articulated and also integrate historical scientific inquiry to improve students' epistemological beliefs about scince in the development and justification of knowledge dimensions.

Conclusion and Suggestion

Learning by doing, experiencing, researching and questioning can change epistemological beliefs (Conley et al., 2004). However, the evidence to date is that inquiry in itself is insufficient to change students' formal epistemologies (Sandoval, 2005). Hence, authentic scientific inquiry that has epistemological differences from simple inquiry was implemented in the physics course and its effects on the students' epistemological beliefs about science were examined in this research.

This study concludes that authentic scientific inquiry experiences help students to enhance their scientific epistemologies through more sophisticated levels. Epistemologically sophisticated students are able and willing to elaborate upon their thoughts and to provide more complex, detailed, and different ideas in their writings and, tend to be more reflective (Liu & Tsai, 2008; Nussbaum, Sinatra & Poliquinet, 2008). Consequently, this study suggests that more studies focusing on students' epistemic sophistication with the help of authentic scientific inquiry experiences would be done.

References

Abd-El-Khalick, F., Boujaoude, S., Duschl, R., Lederman, N. G., Mamlok-Naaman, R., Hofstein, A., ... & Tuan, H. L. (2004). Inquiry in science education: International perspectives. *Science Education, 88*(3), 397-419. https://doi.org/10.1002/sce.10118

Bybee, R. W. (Ed.). (2002). *Learning science and the science of learning: science educators' essay collection*. NSTA press.

Burgin, S. R. (2020) A three-dimensional conceptualization of authentic inquiry-based practices: a reflective tool for science educators, *International Journal of Science Education, 42*(9), 1465-1484, https://doi.org/10.1080/09500693.2020.1766152

Burr, J.E., & Hofer, B.K. (2002). Personal epistemology and theory of mind: Deciphering young children's beliefs about knowledge and knowing. *New Ideas in Psychology, 20*, 199–224.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32–42. https://doi.org/10.3102/0013189X018001032

Chinn, C. A., & Brewer, W. F. (2001). Models of data: A theory of how people evaluate data. *Cognition and Instruction, 19*(3), 323-393

Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education, 86*(2), 175-218.

Conley, A. M., Pintrich, P. R., Vekiri, I., & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology*, *29*(2), 186-204. https://doi.org/10.1016/j.cedpsych.2004.01.004

Hofer, B. K. (2004). Exploring the dimensions of personal epistemology in differing classroom contexts: Student interpretations during the first year of college. *Contemporary Educational Psychology, 29*(2), 129–163. https://doi.org/10.1016/j.cedpsych.2004.01.002

Hofer, B.K., & P.R. Pintrich (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*(1), 88-140.

Liu, C. C., & Tsai, C. C. (2008). An analysis of peer interaction patterns as discoursed by on-line small group problem-solving activity. *Computers & Education, 50*(3), 627-639. https://doi.org/10.1016/j.compedu.2006.07.002

Nussbaum, E. M., Sinatra, G. M., & Poliquin, A. (2008). Role of epistemic beliefs and scientific argumentation in science learning. *International Journal of Science Education, 30*(15), 1977-1999. https://doi.org/10.1080/09500690701545919

Ozkan, Ş. (2008). *Modeling elementary students' science achievement: the interrelationships among epistemological beliefs, learning approaches, and self-regulated learning strategies* [Unpublished doctoral dissertation]. Middle East Technical University.

Rowland, S., Pedwell, R., Lawrie, G., Lovie-Toon, J., & Hung, Y. (2016). Do we need to design course-based undergraduate research experiences for authenticity? *Cell Biology Education, 15*(4), ar79, 1-16. https://doi.org/10.1187/cbe.16-02-0102.

Peffer, M. E., & Ramezani, N. (2019). Assessing epistemological beliefs of experts and novices via practices in authentic science inquiry. *International Journal of STEM Education, 6*, 1-23. https://doi.org/10.1186/s40594-018-0157-9

Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education, 89*, 634–656. https://doi.org/10.1002/sce.20065

Sandoval, W. A., & Morrison, K. (2003). High school students' ideas about theories and theory change after a biological inquiry unit. *Journal of Research in Science Teaching, 40*(4), 369–392. https://doi.org/10.1002/tea.10081

Sandoval, W. A., & Reiser, B. J. (2004). Explanation-driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education, 88*(3), 345-372. https://doi.org/10.1002/sce.10130

Schiefer, J., Golle, J., Tibus, M., Herbein, E., Gindele, V., Trautwein, U., & Oschatz, K. (2020). Effects of an extracurricular science intervention on elementary school children's epistemic beliefs: A randomized controlled trial. *British*

Journal of Educational Psychology, 90(2), 382-402. https://doi.org/10.1111/bjep.12301

Schriebl, D., Müller, A., & Robin, N. (2023). Modelling Authenticity in Science Education. *Science & Education, 32*, 1021–1048. https://doi.org/10.1007/s11191-022-00355-x

Shi, W. Z., Ma, L., & Wang, J. (2020). Effects of Inquiry-Based Teaching on Chinese University Students' Epistemologies about Experimental Physics and Learning Performance. *Journal of Baltic Science Education, 19*(2), 289-297.

Stathopoulou, C., & Vosniadou, S. (2007). Exploring the relationship between physics-related epistemological beliefs and physics understanding. *Contemporary Educational Psychology, 32*, (3), 255-281. https://doi.org/10.1016/j.cedpsych.2005.12.002

Windschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. *Science Education, 92*, 941-967. https://doi.org/10.1002/sce.20259

Wu, H. K., & Wu, C. L. (2011). Exploring the development of fifth graders' practical epistemologies and explanation skills in inquiry-based learning classrooms. *Research in Science Education*, *41*, 319-340. https://doi.org/10.1007/s11165-010-9167-4

Zhao, L., He, W., Liu, X., Tai, K. H., & Hong, J. C. (2021). Exploring the Effects on Fifth Graders' Concept Achievement and Scientific Epistemological Beliefs: Applying the Prediction-Observation-Explanation Inquiry-Based Learning Model in Science Education. *Journal of Baltic Science Education, 20*(4), 664-676.

Augmented Reality Used in Physics Experiments – A New Way to Cognitive Destress Students and Enhance Their Interest?

Hagen Schwanke, Markus Elsholz, Thomas Trefzger

University of Würzburg, Germany

We present a novel approach to visualize scientific concepts in an accessible way to enhance teaching experience in physics education. Dealing with scientific models involves a high degree of complexity for the learners. Both cognitive load due to complexity as well as external load based on the presentation of the content negatively impact learning efficiency. Our objective is to examine if the use of augmented reality can reduce extrinsic load associated with physics lab experiments compared to the use of simulations or classic reply cards. In the augmented reality setting, we overlap digital content to real-life lab experiments in order to improve spatial and temporal contiguity of scientific theory and experiment. In the non-AR settings we provided simulations on a tablet which are used separated from the real-life lab experiment and classic reply cards which are two dimensional paper printings. The applications for the simulation and the AR application "PUMA: Magnetlabor" were specially designed and have already been evaluated in a qualitative usability study.

In a comparative study all participating students experienced all the three presentation options of the supportive learning material. The impact of presentation form on extrinsic load was examined.

Keywords: Augmented Reality (AR), cognitive load, comparative study

Introduction

Experiments continue to be the focus of German science education and scientific research (Lindlahr, 2014; Staatsinstitut für Schulqualität und Bildungsforschung, 2020). In addition, more compulsory student experiments should be carried out in class (Staatsinstitut für Schulqualität und Bildungsforschung, 2020). Due to technical progress, it is possible to use the technology augmented reality (AR for short), which could additionally support the real experiments. However, the effects of AR on learning are still largely unclear and require more detailed investigation (Wyss et al., 2022). For this purpose, an AR application was developed, which will be explored in more detail in this project. In this study 10th grade high school students work on a total of 6 experiments covering aspects of electromagnetism in a 4-hour out-of-school lesson at the university.

Various applications have been developed, which are used by the students while carrying out the experiments. Details of this study are described in the following chapters.

Theory

AR represents a way to enhance a real situation with digital real time information allowing users to move freely in space (Azuma, 1997; Tönnis, 2010). Since the students interact with a technical system, their affinity for technology also plays a major role in that successful interaction with the digital device is an important requirement to perform best on the specific lesson tasks (Attig et al., 2018). Beside other person related issues like subject specific knowledge, interest, and self-concept both the design of the learning environment and the design of the graphical support impact how successful students can master the tasks. This is described by cognitive load theory (Klepsch et al., 2017; R. E. Mayer, 2014; Sweller et al., 1998). Regarding the fact that working memory capacity is limited, the learning environment must be designed not to exceed this total capacity. Cognitive overload (Sweller et al., 1998) describes a situation where students can't access the amount of working memory capacity necessary to implement a learning process. In addition to free capacity, there are three different facets of cognitive load. The intrinsic cognitive load cannot be changed and results from the load caused by the specific content of the topic to be worked on. The germane cognitive load is necessary to understand and memorize the specific content and its connections to other content. This load is conducive to understanding, so that an increase in that load type is desirable. The extrinsic cognitive load depends on the way in which the learning material is presented. By the design of the learning material this load type should be as low as possible. One way to reduce extrinsic load is to increase local and temporal contiguity, i.e. to present dynamic processes which are crucial to understand and carry out the experiment directly on the experimental setup using AR technology. The aim is to prevent cognitive overload by content (intrinsic load) and its graphical representation (extrinsic load), and to free as much capacity as possible for understanding and learning (germane load).

In addition to cognitive load, this article also considers students' physics self-concept. The physics self-concept is an affective construct that is highly relevant to learning processes, as it represents a subjective perception of one's own ability in the subject of physics (Hasselhorn & Gold, 2022). Physics self-concept depends on one's own subject specific performance, and at the same time the performance also depends on one's own self-concept (Köller et al., 2006; Möller & Trautwein, 2009). A connection between self-concept and intrinsic load can therefore be expected because the intrinsic load results from the specific content itself. The self-concept could also have an influence on the extrinsic load, as the visualization of the models via a simulation or AR app is previously unknown

to the students and therefore additional effort is required to understand them. For this reason, the self-concept can also have an influence on how a student individually perceives the representation.

With this theoretical background, the following research question arises:

RQ : *What are the differences in terms of students' cognitive load between the three presentation forms: classic real experiments, experiments enriched with simulations, and experiments equipped with an AR application?*

This article will examine the following hypotheses in detail:

H1: *Students perceived extrinsic cognitive load will differ depending on the presentation form used to provide supportive learning material (AR, simulation, classic form). Extrinsic cognitive load is lowest when the material is presented by augmented reality.*

H2: *There is significant relation between extrinsic cognitive load and affinity for technology. The greater the affinity for technology, the lower the extrinsic cognitive load due to the digital learning material.*

H3: *There is significant relation between extrinsic cognitive load and physics self-concept. The higher the physics self-concept, the lower the extrinsic cognitive load due to the learning material.*

Study Details

Figure 1. Project PUMA



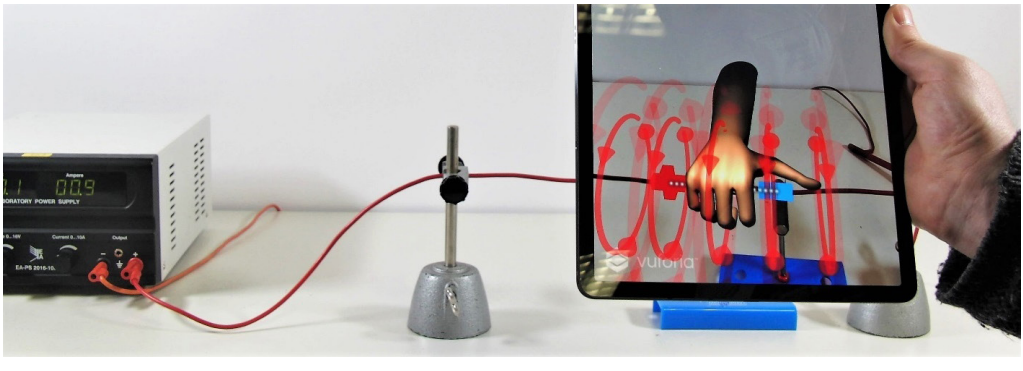
To answer the research question, a laboratory study was carried out. This study is part of the PUMA project. The Physics Lessons with Augmentation project (PUMA for short, see Figure 1.) was initiated at the Chair of *Physik und ihre Didaktik* (Physics Education) at the University of Wuerzburg, Germany with the aim of designing and extensively testing AR applications for various physics topics.

PUMA: Magnetlabor

The application PUMA: Magnetlabor represents a framework application to support students carrying out six different experiments in a *Lehr-Lern-Labor* setting – a student laboratory where students are supervised by pre-service

physics teachers. Being a place where student teachers interact with school students and thereby improve their teaching competencies, the Lehr-Lern-Labor can be titled a teaching lab as well (Elsholz & Trefzger, 2020). All of the experiments base on the E2 experimental kit from Mekruphy (https://mekruphy.com/de/produkte/physik/elektrik-elektronik/experimentiersatz-elektrik-2/). The specific augmentation can be selected for each of the six different stations in the PUMA: Magnetlabor application. The App visualizes microscopic structures and magnetic fields based on established models from the topics of magnetism and electricity, like magnetic field lines or electrons as a bullet. For example, Figure 2 shows the visualization of the concentric magnetic field of a current-carrying conductor.

Figure 2. Visualization of the magnetic field in Oersted's experiment



Using automatically running animations, the internal logic and dynamics of the various processes and their dependencies can be shown in-place in the experimental setup. All AR scenarios allow interactions with the real experiment. In some cases, real-time measurement data is integrated using a Bluetooth connection.

As a second presentation method a simulation application was also created using the same visualizations. The simulation is not locally tied to the experiment and therefore includes the experimental setup only pictographically. Any commercially available tablet can be used to use the two applications. The AR-Apps is available for iOS (https://apps.apple.com/de/app/puma-magnetlabor/id1632458756).

Unity (https://unity.com/) was used to create the two applications. The Vuforia tool provides a reliable tracking method in Unity (https://developer.vuforia.com/), which must also be integrated. The graphic elements and 3D objects were created with Blender (https://www.blender.org/).

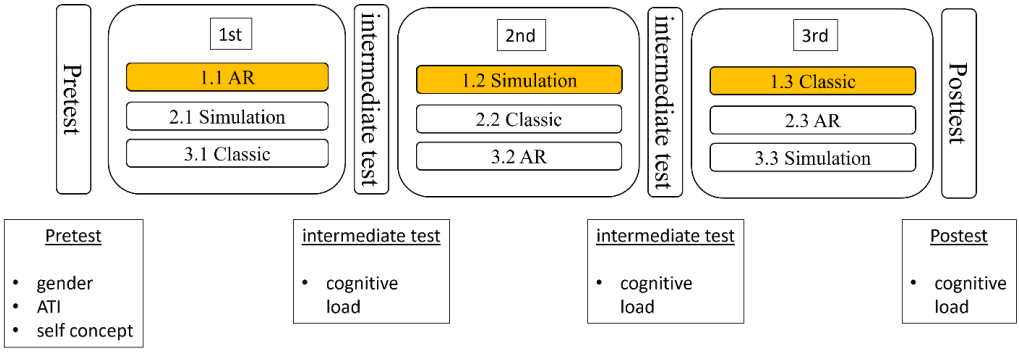
Design and measuring

The focus of the study is the students' individual view of the respective presentation method (classic, AR, simulation) and its influence on the extrinsic cognitive load. The AR- and simulation-Application were described before. The classic presentation method showed some screenshots of the simulation application printed out on paper sheets.

At the beginning of the intervention, the participating students (of a class visiting the student lab) are randomly assigned to three different groups. All students will then receive an iPad to enable paper-free support in all groups and to exclude any possible novelty effect through the use of digital media. In the pretest the physics self-concept scale and the affinity for technology scale are administered. The physics self-concept is measured by a 8-item questionnaire on a 4-point Likert scale based on Habig (2017). The questionnaire for the affinity for technology is based on Franke et al. (2019) and includes 9 items on a 7-point Likert scale between "not true at all" (1) till "completely true" (7). Each student then works on two experiments and one of the three presentation methods of the learning material. Two experiments are considered a one hour working block. Thus, each student carries out six experiments in three working blocks (lasting approx. three hours) and thereby encounters all the three presentation methods. The extrinsic cognitive load is assessed by a questionnaire after each working block. To measure the students' perceived load we used an adapted 9-item questionnaire on a 6-point Likert scale according to Klepsch et al. (2017). This enables us to assign the extrinsic load to the respective presentation method. For this survey an adapted version of the scale published by Klepsch et al. (2017) is used. The test times contains in order to minimize bias due to time effects, the squence in which students encounter the different presentation methods is varied (see Figure 3).

Since December 2022 N=186 10th grade students from several Bavarian high schools in the area of Wuerzburg have taken part (n=88 female, n=98 male) in the study.

Figure 3. Study design and sequence of the intervention and testing



Preparation of the data

All incomplete data sets (n=27) were removed and the internal consistency of the test instruments was determined. An internal consistency of α=0.88 was determined for the technology affinity scale and α=0.92 for the physics self-concept scale. For the subscale of extrinsic cognitive load, there was an internal consistency between α=0.74 and α=0.79 for the three measurement times.

An exploratory factor analysis was carried out for the cognitive load data using a maximum likelihood estimation and Promax rotation. This resulted in the three different factors: intrinsic (ICL), extrinsic (ECL) and germane (GCL) cognitive load (see Table 1).

Table 1. Factor analysis with three assumed factors. Loads below .3 are not shown.

	Cognitive load		
	factor 1	factor 2	factor 3
ICL1		0.96	
ICL2		0.62	
ICL3		0.40	
ECL1	0.78		
ECL2	0.75		
ECL3	0.68		
GCL1			0.60
GCL2			0.87
GCL3			0.41

Preliminary results

To address hypothesis H1 we analyzed the differences in means of extrinsic cognitive load (ECL) due to the different presentation methods (AR: augmented reality, SIM: simulation, K: classic presentation) using a confirmatory factor analysis model. According to Hu and Bentler (1999) the model provides a very good fit to the data: CFI=0.987, RMSEA=0.042 and SRMR=0.044. For that model we found the values for ECL mean and standard deviation (sd) shown in table 2:

Table 2. mean and standard deviation for the extrinsic cognitive load

	mean	sd
AR_ECL	2.65	0.98
SIM_ECL	2.78	1.03
K_ECL	2.62	0.97

There are no significant differences in the perceived extrinsic cognitive load between the three different presentation methods, thus H1 has to be rejected. In a next step, we fitted a strong invariant multi-group model with respect to gender, see table 3 for the respective ECL mean values:

Table 3. mean and standard deviation for the extrinsic cognitive load with respect to gender

	male		female	
	mean	sd	mean	sd
AR_ECL	2.64	1.09	2.66	0.86
SIM_ECL	2.69	1.04	2.86	1.05
K_ECL	2.57	0.96	2.68	0.99

As the mean values do not differ significantly due to gender we conclude that there is no relevant impact of gender on extrinsic cognitive load in the setting of this study.

Table 4 shows (manifest) mean values and standard deviations for the affinity for technology (ATI) scale and the physics self-concept (SC) scale with respect to gender.

Table 4. mean and standard deviation for female and male

	male		female	
	mean	sd	mean	sd
ATI	3.98	0.82	3.45	0.80
SC	2.87	0.63	2.55	0.56

There are significant differences between the two gender groups in both ATI (t(161) = 4.0958, p = 6.647e-05) and SC (t(161) = 3.3921, p = 8.727 e-04) .

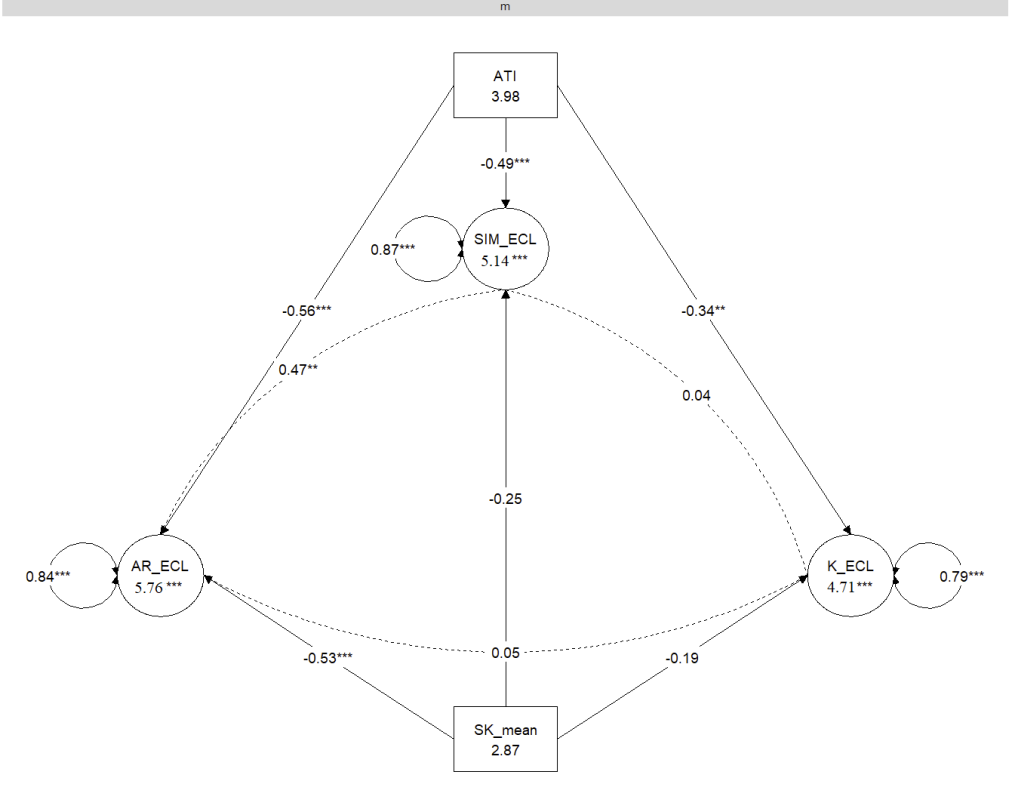
A multi-group structural equation analysis was conducted to examine the relationships between extrinsic cognitive load, affinity for technology (H2) and self-concept (H3) for female and male participants respectively. For a better overview, the two groups are first examined individually followed by a general discussion. These notes apply to both Figure 4 and Figure 5:

- p < .05; ** p < .01; *** p < .001;
- the coefficients on solid lines correspond to standardized path coefficients β.

Results and Discussion

Male Students

Figure 4. Result of the structural equation model of the male students



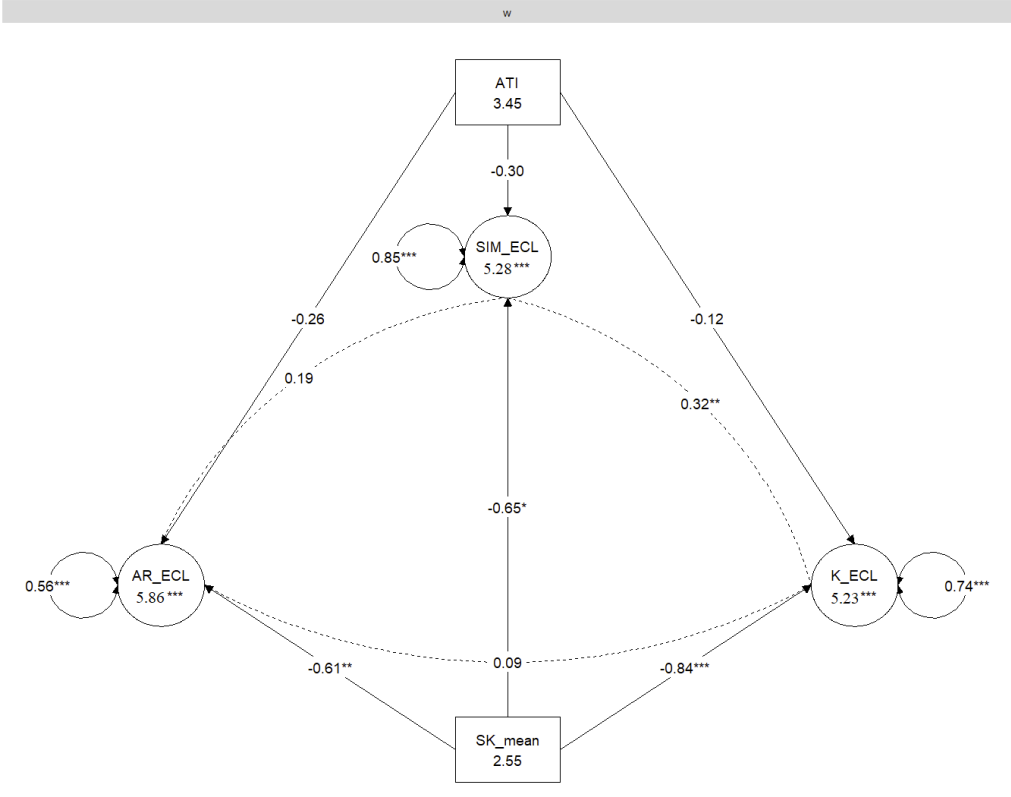
As can be seen, there is low correlation between K_ECL and both AR_ECL (0.05) and SIM_ECL (0.04). This is consistent with the assumption that there is no systematic relationship between classic and digitally enhanced forms of presentation. The relation between AR_ECL and SIM_ECL (0.47**) underlines this statement, as similar visualizations were used for both digital applications.

It can also be stated that technology affinity is a stronger predictor for extrinsic load when dynamic digital representation methods are used (ATI ~ AR_ECL: -0.56***; ATI ~ SIM_ECL: -0.49***) compared to the classic presentation method (ATI ~ K_ECL: -0.34**). This is consistent with the theory of cognitive load, as it can be assumed that a higher level of technology affinity leads to lower extrinsic cognitive load. This could be due, among other things, to the fact that male students who are not technically affine are more quickly overwhelmed by problems with AR or simulation applications because they are less familiar with solution strategies and therefore find the learning material more cognitively stressful. However, as there is a significant relation between ATI and ECL for all presentation methods, H2 can be accepted for the group of male participants.

Physics self-concept strongly predicts extrinsic cognitive load when the learning material is presented by the AR App (SK_mean ~ AR_ECL: -0.53***). This pronounced relationship of self-concept with extrinsic load may be a sign of the effort necessary to understand the presented models. The visualization of the AR application requires a high level of abstraction and that this ability had not been trained before. Furthermore, it can be seen that there is a small but not significant connection between the self-concept and the simulation and the classic representation (SK_mean ~ SIM_ECL: -0.25; SK_mean ~ K_ECL: -0.19). In general, it can be said that extrinsic cognitive load in all representation methods is negatively correlated with self-concept, but since only one of the path coefficients is significant, H3 can only be accepted in the case of the AR presentation form. In order to be able to make a more detailed statement maybe additional constructs, e. g. a performance measure, should be included in the analyses.

Female students

Figure 5. Result of the structural equation model of the female students



There is no substantial relationship between extrinsic cognitive load in the AR and classical representation method (K_ECL ~ AR_ECL: 0.09). Although the same visualization models used in the AR and the simulation visualization, there is very low correlation between the respective ECL measures (AR_ECL ~ SIM_ECL: 0.19). Somewhat unexpected is the substantial relationship between the cognitive load in the classic and the simulation representation method (K_ECL ~ SIM_ECL: 0.32**). This contradicts the assumption that there is no connection between dynamic (simulation) and static (classic) visualization with respect to cognitive load.

The relationship between cognitive load and technology affinity is much more in line with theory. There are greater values for the AR (-0.26) and simulation (-0.30) representation methods than for the classical one (-0.12). Because all path coefficients are not significantly different from zero, H2 cannot be confirmed for female students.

The physics self-concept, on the other hand, has a significant relationship on

the extrinsic load of the respective presentation forms (SK_mean ~ AR_ECL: -0.61**; SK_mean ~ SIM_ECL: -0.65*; SK_mean ~ K_ECL: -0.84***). In this case, H3 can be confirmed for female students. The fact that the self-concept is strongest related to the classical presentation form could be, among other things, due to the fact that this type of visualization already had been known from previous lessons and therefore only a low level of abstraction skills is required.

Discussion and limitations

The use of augmented reality in terms of supporting students does not reduce students perceived extrinsic load compared to other methods of presenting supportive material, i. e. providing simulations or static representations on a digital device. Actually, it was to be expected that the extrinsic load would be lowest in the augmented reality presentation form due to high spatial and temporal contiguity. A possible explanation could be that the attending students used this technology for the first time and no detailed training was given, which means that it maybe was too challenging to handle. In addition, students used the AR App approximately 15 minutes during the one-hour experimental block. Maybe this short integration of the AR application could not have been precisely delineated by the test.

Analyses of relations between extrinsic cognitive load and affinity for technology and physics self-concept yield somewhat inconsistent results with respect to gender. As technology affinity is related with the extrinsic cognitive load among male students (in an expected manner, i.e. stronger relation in the AR presentation form than in simulation or classic presentation forms), it is not among female students. Regarding physics self-concept it is exactly the other way around, i.e. as self-concept is significantly related with extrinsic cognitive load among female students, it isn't in the group of male students. An interpretation approach might be that male students tend to focus on technical aspects of the supportive learning material whereas female students rather focus on the content visualized.

It must be pointed out that at the moment a more detailed data analysis (e. g. additional groups) is limited by sample size (N=159) For this reason, further data collection is already in progress so that further analyses will shed light on the different patterns of relations between extrinsic load and predictor variables among male and female students.

Acknowledgment

This project is funded by the Universitätsbund Würzburg and Deutsche Telekom Stiftung.

References

Attig, C., Mach, S., Wessel, D., Franke, T., Schmalfuß, F., & Krems, J. (2018). Technikaffinität als Ressource für die Arbeit in Industrie 4.0. In Professur Arbeitswissenschaft und Innovationsmanagment (Chair), *Interact Conference.: Innovation der Innovation - Neu Gedacht, Neu Gemacht.* Symposium conducted at the meeting of Technische Universität Chemnitz, Chemnitz.

Azuma, R. T. (1997). A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments*, *6*(4), 355–385. https://doi.org/10.1162/pres.1997.6.4.355

Elsholz, M., & Trefzger, T. (2020). Identitätsarbeit im Lehr-Lern-Labor-Seminar: Dimensionalität und Veränderung des akademischen Selbstkonzepts angehender Physiklehrkräfte während einer zentralen Praxisphase. In D. Bosse, M. Meier, T. Trefzger, & K. Ziepprecht (Eds.), *Lehrerbildung auf dem Prüfstand: 1-2020. Professionalisierung durch Lehr-Lern-Labore in der Lehrerausbildung* (pp. 78–100). Empirische Pädagogik.

Franke, T., Attig, C., & Wessel, D. (2019). A Personal Resource for Technology Interaction: Development and Validation of the Affinity for Technology Interaction (ATI) Scale. *International Journal of Human–Computer Interaction*, *35*(6), 456–467. https://doi.org/10.1080/10447318.2018.1456150

Habig, S. (2017). *Systematisch Variierte Kontextaufgaben und Ihr Einfluss Auf Kognitive und Affektive Schülerfaktoren*. *Studien Zum Physik- und Chemielernen Ser: v.223*. Logos Verlag Berlin. https://ebookcentral.proquest.com/lib/gbv/detail.action?docID=5313492

Hasselhorn, M., & Gold, A. (2022). *Pädagogische Psychologie: Erfolgreiches Lernen und Lehren* (5. überarbeitete Auflage). Kohlhammer Verlag. https://doi.org/10.17433/978-3-17-039783-5

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and Validation of Two Instruments Measuring Intrinsic, Extraneous, and Germane Cognitive Load. *Frontiers in Psychology*, *8*, 1997. https://doi.org/10.3389/fpsyg.2017.01997

Köller, O., Trautwein, U., Lüdtke, O., & Baumert, J. (2006). Zum Zusammenspiel von schulischer Leistung, Selbstkonzept und Interesse in der gymnasialen Oberstufe. *Zeitschrift Für Pädagogische Psychologie*, *20*(1/2), 27–39. https://doi.org/10.1024/1010-0652.20.12.27

Lindlahr, W. (2014). Virtual-Reality-Experimente für Interaktive Tafeln und Tablets. In J. Maxton-Küchenmeister & J. Meßinger-Koppelt (Eds.), *Naturwissenschaften. Digitale Medien im naturwissenschaftlichen Unterricht* (pp. 90–97). Joachim-Herz-Stiftung Verlag.

Mayer, R. E. (2014). Cognitive Theory of Multimedia Learning. In R. Mayer & R. E. Mayer (Eds.), *Cambridge handbooks in psychology. The Cambridge handbook of multimedia learning* (Second edition, pp. 43–71). Cambridge University Press.

https://doi.org/10.1017/CBO9781139547369.005

Möller, J., & Trautwein, U. (2009). Selbstkonzept. In E. Wild & J. Möller (Eds.), *Springer-Lehrbuch. Pädagogische Psychologie: Mit 27 Tabellen* (pp. 179–203). Springer. https://doi.org/10.1007/978-3-540-88573-3_8

Schneider, R., Gentrup, S., Jansen, M., & Stanat, P. (2022). Kohortentrends in schulfachbezogenen Selbstkonzepten und Interessen bei Mädchen und Jungen. *Zeitschrift Für Pädagogische Psychologie.* Advance online publication. https://doi.org/10.1024/1010-0652/a000346

Staatsinstitut für Schulqualität und Bildungsforschung. (2020). *LehrplanPLUS: Fachlehrplan - Gymnasium Physik*. https://www.lehrplanplus.bayern.de/schulart/gymnasium/fach/physik/inhalt/fachlehrplaene

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, *10*(3), 251–296. https://doi.org/10.1023/A:1022193728205

Tönnis, M. (Ed.). (2010). *Informatik im Fokus*. *Augmented Reality*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-14179-9

Wyss, C., Degonda, A., Bührer, W., & Furrer, F. (2022). The Impact of Student Characteristics for Working with AR Technologies in Higher Education—Findings from an Exploratory Study with Microsoft HoloLens. *Information*, *13*(3), 112. https://doi.org/10.3390/info13030112.

Researching the Teaching and Learning of Physics in Formal and Informal Contexts: A Study from Kindergarten to University

Giancarlo Artiano[1,2], Emilio Balzano[3]

[1]Università degli Studi della Campania Luigi Vanvitelli, Italy

[2]Università degli Studi Suor Orsola Benincasa, Italy

[3]Università degli Studi di Napoli Federico II, Italy

Modeling physical phenomena and acquiring mathematical formalisation can be challenging in physics education at all levels. Therefore, it is recommended that interpretative and mathematical models be developed gradually from an early age, starting from kindergarten. Even young children possess innate abilities of abstraction, discerning discrete quantities, and spatial thinking, making it possible to introduce them to the basics of modeling physical phenomena. Our research is based on a cognitive dynamic model derived from years of direct observation and research activities in real-life learning and teaching contexts. Our cognitive model is based on the metaphorical notion of resonance. This suggests that people understand the meaning of physical and mathematical theories by resonating with the formal models of reality presented by physics. In this regard, we present a phenomenological approach to modeling physics phenomena that focuses on developing didactic experimentation from kindergarten to university courses. The didactic experimentations we designed and implemented in teacher training courses and classrooms are based on a revisited approach to Design Experiment. The research results are based on action-research activities implemented in formal and informal contexts and analysed qualitatively using a content analysis approach.

Keywords: Physics, modeling, resonance

Introduction

What kind of physics should be taught to enable people to reconstruct facts? This is one of the main questions that has driven our research study for several years. Researching ways of learning and teaching physics is a complex process of reality (Arcà, & Guidoni,1987) and requires transversal knowledge in different but interconnected cultural fields (Cini,1991). In the cultural framework, physics plays a fundamental role in the relationship with history and the other sciences, whether natural or social sciences or humanities (Cini et al., 1971). Examining different ways how people learn physics concepts directly impacts

improving critical thinking in humans of all ages (Viennot, & Décamp,2020). Unlike other natural sciences, physics is interested in things that exist and interact with us, in time and space. Any understanding of the world requires a careful selection of variables and an appropriate language to describe their relationships: mathematics. When a fact or phenomenon is observed to acquire more knowledge, the whole person, with their experience, previous knowledge (related to the culture in which they live) and perceptions, are involved in the observation. With the first observations of the phenomenon, the person develops several mental images that do not immediately correspond to the images of the natural world (Lehrer & Schäuble, 2015; Odden & Russ, 2019). To understand the phenomenon according to the formal models of physics, it is necessary to become familiar with its language consisting of schematisation and mathematical tools, which are useful for operating on abstract entities corresponding to schematised reality (Upmeier zu Belzen et al., 2019). Identifying the cause-and-effect relationship between the variables of a particular phenomenon allows people to be fully aware of the meaning of the model. If all the complexity of the cognitive organisation of facts is lost, the sciences, particularly physics, lose their cultural significance and social value. In 2021, the Nobel Prize for Physics Giorgio Parisi raised a severe alarm about the relationship between science and society in Italy. He states that

> there are anti-scientific solid tendencies in today's society. The prestige of Science and trust in it are decreasing rapidly. This mass distrust in science that reaches our parliament may also be due to a certain arrogance of those scientists who present science as absolute wisdom and [...] the slow decline of public school.

This statement is confirmed by national surveys of science learning at all levels of education (Report Invalsi 2022). The problem is complex, and the current difficulties have deeper origins, which must be fully understood to counter them. We believe that we need to go beyond quantitative data and work to develop models of interventions that can respond to people's specific needs. Over the years, our research group has gathered many experiences (Balzano, 2005; Balzano et al., 2014; Artiano & Balzano, 2023). We operate in our territory through training courses that offer physics educational activities in formal and informal contexts for many kinds of people: from children to adults, from students to teachers, educators, and families. These experiences are documented on the community website, which today welcomes a broad educational audience (623 people) made up of students, researchers, teachers, and educators. The research focuses on the cognitive resonance model (Guidoni et al. 2005a). It is a theoretical framework that helps us interpret human knowledge. It is also a valuable tool for didactic design experiments (e.g., Cobb et al., 2003; diSessa & Cobb, 2004). Therefore, our teaching experiments follow a phenomenological approach to

the modeling process (Lehrer & Schäuble, 2015; Schwarz et al., 2009). The methods of investigation that we have used include observation, documentation, interviews, and focus groups. We have analysed the data qualitatively using a content analysis approach (Hsieh & Shannon, 2005; Lichtman, 2013; Mayring, 2014). In the following paragraphs, we will describe our research in physics education and the results we have achieved in more detail.

Methods

The process of teaching and learning, as well as the understanding and representation of knowledge, is a significant global challenge today. For centuries, theologians, philosophers, and scientists have attempted to comprehend the nature of human knowledge, resulting in many theories on the subject. Modern data collection techniques have allowed it to delve deeper into the brain's workings during cognitive activities in humans and other animals (Dehaene,2009; Dehaene, 2011; Rizzolatti & Sinigaglia, 2023). These studies often support previously developed psychological theories (Piaget,1975; Vygotsky 1964) or common-sense intuitions while disproving other theoretical assumptions. Scientific disciplines, such as physics, appear abstract and may seem suitable to be taught and learned only at certain stages of development (Piaget,1975; NGSS Lead States; 2013). However, the challenge in the dynamics of learning and teaching must be approached through research that connects the psychological structure of the students and their socio-cultural background with the logical and scientific structure of the subject matter. This requires us to ask two key questions: 1) What are the fundamental structures of a discipline that the mental structures of a particular age group can comprehend? 2) How can we encourage the development of these structures? To address these questions, our research group has developed a cognitive model that can help individuals better understand the fundamentals of physics by resonating with their ways of representing abstract mental structures.

The Resonance Cognitive Model

Over the years, our research group has worked to establish a general cognitive framework for revising disciplinary contents of all levels of education. The Resonance Cognitive Model is the cornerstone of our scientific activities (Guidoni et al. 2005a). It is based on the concept of resonance. We can consider the case of resonance in the swing to get an idea of the model. The oscillating period of the swing and the timed pushes must be in resonance to keep the swing going very efficiently. In the same way, suitable didactic mediation occurs when people's interpretations of reality are in resonance with the formal models that physics has of reality. In other words, we can say that learning physics is practical when people are helped in the modelling process. This is possible by considering individual cognitive aspects and the socio-cultural background in

designing scientific activities. People who follow our activities are involved both in direct laboratory experiences and in the modeling of the phenomena under study. During activities, we support the cognitive reconstruction of the facts by asking our targets to share continuous rethinking of their own way of modeling. This way, we try to critically engage people with metacognitive strategies to consolidate their learning (Arcà, & Guidoni,1987).

The Revisiting Of The Contents

Our research group has been working on developing a model of cognitive dynamics. In addition, we have also revisited the contents of the physics curriculum to facilitate the understanding of the discipline. Our work began with a critical analysis of the national guidelines for the Italian Ministry of Education curriculum (2012), and a study of the relevant literature (NRC,2012). Our experimentation and training activities, present physics to students, teachers, and trainees by exploring, modeling, and mathematising physical phenomena. We focus on two interrelated areas: *crosscutting concepts* (or ways of looking) and *key concepts* (which pertain to the different areas of the phenomenology of the discipline). The *crosscutting concepts* we identified are: - system and interaction between systems; - variables and representation; - measure; - equilibrium, conservation, and invariance. *Key concepts* of physics include: - the structure and properties of matter and materials, - force; - equilibrium and motion; - gravitational force; - field; - energy; - the relationship between force and energy; - heat and temperature; - the propagation of light; - waves and oscillations; - electricity and magnetism. We develop activities on two interrelated levels, which include understanding the *key concepts* of physics and transversal concepts that apply to different scientific disciplines. We also encourage continuous reflection on the overall development of students, using strategies that emphasise how knowledge and skills mature by working simultaneously on physics and language, physics and mathematics, and physics and technology.

Scientific Activities

Our research endeavours encompassed a comprehensive analysis across various real teaching-learning contexts focusing on revisiting the Design Experiment approach (Cobb et al, 2003; Di Sessa & Cobb, 2004). From our experiences, we have seen a tendency to simplify the teaching of phenomena by presenting ideal situations at all levels of education (i.e., an inclined plane without friction or a simple harmonic oscillator). When educators solely rely on showcasing perfect or ideal scenarios to students, they inadvertently hinder the development of the students' innate ability to comprehend and navigate reality through intricate cognitive frameworks. Such an approach to teaching can lead to a disconnection between what is being taught and what students actually experience in the real world. Therefore, it is crucial for educators to present a balanced perspective

that encompasses both ideal and realistic scenarios to promote a comprehensive understanding of the subject matter. Activate the body, emotions, and imagination through laboratory experiences (both qualitative and quantitative) to help students connect physical modeling and mathematical understanding of phenomena. By way of example, we report below some of our scientific activities that represent both our vertical (from kindergarten to university) and horizontal (in formal and informal contexts) educational action.

An intuitive approach to the dynamic systems

In recent years, we have experimented with an approach to studying the laws of evolution of dynamics based on mathematical discretization (Amabile et al., 2023). This approach is introduced as early as middle school and is effective up until university studies. The purpose is to enable students to predict the evolution of natural physical systems using the most appropriate mathematical tools for their level of education. We believe it is essential for students to have a clear understanding of the formal apparatus of a physical theory and to be able to master computational methods to verify the effectiveness of the theory. This approach allows us to deal with cases of great physical interest, sometimes not present in textbooks and programs because they require more advanced mathematics. The examples presented range from studying oscillations, waves, and gravitational systems.

Training teachers

The ability to teach is not innate; even if there is passion, motivation, and authentic communication, it cannot be improvised. It requires a long process of initial and ongoing training and interactions with other teachers, educators, and researchers. Since 2018, we have endeavoured to provide a national training program that caters to approximately one hundred participants (Amabile et al., 2022). This program is inclusive of teachers from kindergarten through middle school (grades K-8), as well as pre-service teachers. Our primary objective is to encourage pre-and in-service teachers to share their didactic materials and exemplary experiences, utilising an action-research approach that transcends the confines of a purely academic setting. Our approach is designed to provide teachers with the tools and resources they need to succeed in their classrooms. Teachers can collaborate and learn from one another through the sharing of didactic materials and exemplary experiences, thereby fostering a sense of community and shared purpose. Teachers engage in course activities and incorporate research-based suggestions and methods into their teaching. The suggestions in this regard pertain to how physics can contribute to designing educational paths that develop skills in teaching and learning related to science, language, mathematics, physics, and technology.

Science education for low-income communities

We have gained significant experience in local and international science education projects over the years. Our group has been contributing to scientific education's role in citizens' cultural backgrounds, and this requires sharing a non-traditional approach to teaching science to the educational community. Recently, our focus has been on the specific role that science education can play in high-risk areas. In such areas, providing quality education in Science, Technology, Engineering, Art, and Math (STEAM) disciplines can be a challenge. Our interventions have focused on pupils' parents, aiming to extend the exploration of scientific phenomena to the home, where children can experiment and revisit their experiences at school. Our researchers work with motivated parents who significantly contribute to studying their children at home by supporting reasoning and encouraging the development of critical thinking. Parents are involved in scientific activities in specific laboratories, allowing them to discover the vast wealth of technical and scientific knowledge that enables us to live but which we cannot communicate in the way that misinterpreted accredited science would like. We have started a participatory planning process, which involved the school "I.C. Porchiano-Bordiga" of Naples and the social enterprise "Verde Speranza", on the construction of a pond. The pond was built in a former school complex, which today houses numerous third-sector associations and will be the place where local school pupils and their families can carry out numerous experimental scientific activities. The activities are based on analysing systems and ecosystems, possibly measuring physical variables, such as temperature, humidity and dissolved oxygens, with online transducers.

Findings and Discussion

Our approach to tackling the challenges of science education was formulated through a collaborative effort that involved the cooperation of students, teachers, researchers, educators and families. Our approach was predicated upon understanding, intervention, research, and observation. The two primary social activities involved in this undertaking, namely the practice of science and the dissemination of scientific knowledge, were pursued as collective endeavours with significant implications for individuals and society.

We have garnered significant knowledge from our research experiences in real-life teaching-learning contexts, ranging from teaching university courses to conducting teacher training and engaging with at-risk children and their families. All the didactic material we have collected thus far can be accessed via peer-reviewed papers and our Research Group website (www.les.unina.it). Our multimedia documentation (video, photo and elaborations) of the activities undertaken is intended to foster a connection between teachers, researchers, educators, and students. It also encourages the latter to review their scientific

experiences at home with their families cognitively.

After many years of direct involvement, the results obtained thus far have been very promising by all parties concerned. Effective intervention models that can address the needs of the territories we operate are becoming evident. What have we learned from these experiences? Firstly, we have comprehended that practical proposals for science education can only be developed through a profound integration effort between experts in scientific disciplines, mathematics, educators, psychologists, teachers, and families. The engagement of the education community represents the principal long-term strategy and is indispensable for the deliberate dissemination of science education on a large scale. Secondly, despite the data indicating a lack of scientific knowledge, we have observed that the emergence of scientific dialogue is feasible in all socio-cultural contexts if additional support is provided, such as mediators employing everyday language, local cultural experiences, and a phenomenological approach to studying scientific disciplines. This method of relating to people through resonance is efficacious because it heightens their self-assurance and helps them to comprehend their knowledge. Physics education researchers need to interact with people and communicate the meaning of doing science a daunting challenge fraught with inevitable uncertainties but with considerable rewards.

References

Amabile, A., Annunziata, A., Artiano, G., & Balzano, E. (2022). Experimentation and Research in the Physics Course for the Preparation of Primary School Teachers in Naples. *Education Sciences, 12(4),* 241.

Amabile, A., Annunziata, A., Balzano, E., Piccialli, P., Figari, R., & Artiano, G. (2023). An intuitive introduction to the evolution of physical systems. *Il Nuovo Cimento, 100(198),* 46.

Arcà, M., & Guidoni, P. (1987). *Guardare per sistemi, guardare per variabili*. AIF Editore.

Artiano, G., & Balzano, E. (2023). Teaching and learning physics in formal and informal contexts. *Il Nuovo Cimento, 100(182),* 46.

Balzano, E., de Ambrosis, A., Gagliardi, M., Giordano, E., Guidoni, P., Mendella, G., & Tarsitani, C. (2005). A research on the conceptual organization of physics'curriculum and standards. *Proceedings of Durban ICPE2004, What Physics Should We Teach,* 107-115.

Balzano, E., Cuomo, F., Minichini, C., & Serpico, M. (2014). Communities of practice and continuous teacher professional development. Findings from eight case studies. In *Proceedings of the ESERA 2013 conference: science education research for evidence-based teaching and coherence in learning. Atenas.*

Chiu, M.-H., & Lin, J.-W. (2019). Modelling competence in science education. *Disciplinary and Interdisciplinary Science Education Research,* 1, 1–11.

Cini, M., Ciccotti, G., De Maria, M., & JonaLasinio, G. (1971). *L'ape e l'architetto.*

Cini, M. (1991). Science and sustainable society. *Capitalism Nature Socialism, 2(2)*, 39-54.

Cobb, P., Confrey, J., diSessa, A.A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9-13.

Dehaene, S. (2009). Origins of mathematical intuitions: The case of arithmetic. *Annuals of the New York Academy of Sciences, 1156(1),* 232-259.

Dehaene, S. (2011). *The number sense: How the mind creates mathematics*. OUP USA.

DiSessa, A.A., & Cobb, P. (2004). Ontological innovation and the role of theory in design experiments. *Journal of the Learning Sciences, 13(1),* 77-103.

Guidoni, P., Iannece, D., & Tortora, R. (2005). Forming Teachers as Resonance Mediators. *International Group for the Psychology of Mathematics Education, 3,* 73-80.

Hsieh, H.F., & Shannon, S.E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*(9), 1277-1288.

Kenyon, L., Schwarz, C., & Hug, B. (2008). The benefits of scientific modeling. *Science and Children, 46*(2), 40-44.

Lehrer, R., & Schauble, L. (2015). The development of scientific thinking. In L. S. Liben, U. Müller, & R.M. Lerner (Eds.), *Handbook of child psychology and developmental science: Cognitive processes* (p. 671–714). Wiley.

Lichtman, M. (2013). Qualitative research in education: A user's guide (3rd ed.). *SAGE Publications.*

Mayring, P. (2014). *Qualitative content analysis: theoretical foundation, basic procedures and software solution*. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-395173.

National Research Council (NRC). (2012). *A Framework for K-12 Science Education*. National Academies Press.

NGSS Lead States (2013). *Next generation science standards: For states, by states*. National Academy Press.

Nicolaou, C.T., & Constantinou, C.P. (2014). Assessment of the modeling competence: A systematic review and synthesis of empirical research. *Educational Research Review, 13*, 52–73.

Nielsen, S.S., & Nielsen, J.A. (2021). A competence-oriented approach to models and modelling in lower secondary science education: Practices and rationales among Danish teachers. *Research in Science Education, 51*, 565-593.

Odden T.O.B. & Russ R.S. (2019). Defining sensemaking: Bringing clarity to a fragmented theoretical construct.*Science Education, 103(1),* 187-205.

Passmore, C., Stewart, J., & Cartier, J. (2009). Model-based inquiry and school science: Creating connections. *School Science and Mathematics*, 109(7), 394-402.

Piaget, J. (1975). Comments on mathematical education. *Contemporary education, 47(1),* 5.

Rizzolatti, G., & Sinigaglia, C. (2023). *Mirroring Brains: How we understand others*

from the inside. Oxford University Press.

Schwarz, C.V., Reiser, B., Davis, E., Kenyon, L., Acher, A., Fortus, D., Shwartz, Y., Hug, B., & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching, 46(6)*, 632-654.

Siraj-Blatchford, I., Sylva, K., Muttock, S., Gilden, R., & Bell, D. (2002). *Researching Effective Pedagogy in the Early Years (REPEY) DfES Research Report 365*. Queen's Printer.

Upmeier zu Belzen, A., van Driel, J.H., & Krüger, D. (2019). Introducing a framework for modeling competence. In A. Upmeier zu Belzen, D. Krüger, & J.H. van Driel (Eds.), *Towards a competence-based view on models and modeling in science education* (p. 3-19). Springer International Publishing.

Viennot, L., & Décamp, N. (2020). *Developing critical thinking in physics.* Springer International Publishing.

Vygotsky, L. S. (1964). *Thought and language.*



Strand 4

Digital Resources for Science Teaching and Learning

María Napal & Leticia Garcia-Romano

Strand Chairs/Co-editors

Foreword

This chapter, titled Digital Resources for Science Teaching and Learning, is part of the ESERA 2023 Conference Proceedings and brings together 15 studies that explore the rapidly evolving role of digital tools and technologies in science education. As digital resources continue to transform educational landscapes, their effective integration into science teaching and learning has become a critical area of focus for educators, researchers, and policymakers.

The proceedings in this chapter highlight a range of innovative approaches that leverage digital resources to enhance science learning experiences. From interactive simulations and virtual labs to online platforms and multimedia resources, the studies present diverse examples of how technology can support deeper understanding, improve engagement, and offer personalized learning pathways. These contributions underscore the potential of digital tools not only to convey scientific concepts but also to cultivate inquiry, critical thinking, and collaborative learning.

In addition to exploring the benefits of digital resources, the research in this chapter addresses challenges such as equitable access, effective implementation, and teacher preparedness. The studies consider the importance of professional development in equipping educators with the skills and confidence needed to integrate digital tools meaningfully into their teaching practices. Furthermore, the chapter explores how digital resources can be designed and utilized to foster inclusive learning environments that cater to diverse learners.

The collective insights presented in this chapter are timely and relevant as science education increasingly relies on technology to reach and engage students. These proceedings provide valuable perspectives on both the opportunities and challenges of digitalization in education, offering evidence-based practices and strategies for maximizing the impact of digital resources.

We extend our sincere appreciation to the authors and reviewers who contributed to this chapter. Their work offers a glimpse into the future of science education, where digital resources play a central role in shaping how students learn and interact with scientific knowledge.

As you read through the studies in this chapter, we hope you find inspiration and practical insights that will inform your own work in integrating digital resources into science teaching and learning.

Teaching Nanotechnology Concepts Through the Use of Digital Technology to Early-Primary School Children

Pandora Dorouka[1], Michail Kalogiannakis[2]

[1]Department of Preschool Education, University of Crete, Greece

[2]Department of Special Education, University of Thessaly, Greece

This research has the objective to evaluate the possibilities offered in order to improve the existing teaching procedures through computer desktops and smart mobile devices. Nanoscience and Nanotechnology (NST) education is a new interdisciplinary field that promises to address global challenges and it could immensely benefit from the applications of technology. It is and it should be viewed as an urgent priority so that the numbers of specialized human resources rise to the point that are capable of sustaining a solid and prosperous NST-related industry. In order for that to be achieved, it is important to popularize the existing NST concepts to the younger generations and specifically to young children. The focus of this research is to compare the impact and effectiveness of alternative experiential teaching on early-primary school children of the domain of the interdisciplinary field of NST against its more contemporary alternatives that employ the use of computers and smart mobile devices. Through the use of software applications in the form of digital games, the context of this research is to expose groups of young children to elements of NST concepts and examine the learning outcomes. The three-step research process aimed to measure the effectiveness of two different digital technologies (computers and tablets) at an introductory level on children's understanding of Nanotechnology concepts. 150 second-grade children were divided into two experimental groups and a control group. Children's knowledge about size was assessed with the Nanotechnology Elementary Knowledge Assessment Test (TENANO). The findings revealed that the two experimental groups significantly outperformed the post-test control group, with the dominant tablet group.

Keywords: Early childhood, nanotechnology, mobile learning.

Introduction

Mobile devices have become an integral part of our daily lives with advances in technology and modern developments. Researchers, educators, and parents pay special attention to children's digital education. The critical factors for digital

learning include content, methods, and techniques for using digital technologies, which make the content more exciting and engaging. Despite the proliferation of research focusing on learning through digital technologies in preschool and primary education, as evidenced by recent literature, there is a significant lack of developmentally appropriate digital apps that educationally encourage early childhood children's interaction with nanotechnology; although children have experienced the unprecedented health crisis with the new coronavirus (80-120 nm) (Zhou, 2022), a distinct nanoscale entity (Delgado et al., 2015).

Tablets change into the preferred device for young children due to their big enough screen size, high mobility, offline and online media streaming ability, low cost and advanced interactive capabilities. These devices are appropriate for the young children since they are very mobile and do not require particular motor skills, as they do not need the users to utilize a keyboard or a mouse (Kucirkova et al., 2014and using the concept of exploratory talk as developed by Mercer et al. to analyse peer engagement. For both approaches, quantitative and qualitative indices of children's engagement were considered. The overall findings suggested that in terms of the Bangert-Drowns and Pyke taxonomy, the quality of children's individual engagement was higher with the OS app in contrast to their engagement with other app software. The frequency of children's use of exploratory talk was similar with the OS and colouring and drawing apps, and a detailed qualitative analysis of the interaction transcripts revealed several instances of the OS and drawing apps supporting joint problem-solving and collaborative engagement. We suggest that critical indices of an app's educational value are the extent to which the app supports opportunities for open-ended content and children's independent use of increasingly difficult features. © 2013 Elsevier Ltd. All rights reserved.,"author":[{"dropping-particle":"","family":"Kucirkova","given":"Natalia","non-dropping-particle":"","parse-names":false,"suffix":""},{"dropping-particle":"","family":"Messer","given":"David","non-dropping-particle":"","parse-names":false,"suffix":""},{"dropping-particle":"","family":"Sheehy","given":"Kieron","non-dropping-particle":"","parse-names":false,"suffix":""},{"dropping-particle":"","family":"Fernández Panadero","given":"Carmen","non-dropping-particle":"","parse-names":false,"suffix":""}],"container-title":"Computers and Education","id":"ITEM-1","issued":{"date-parts":[["2014"]]},"page":"175-184","publisher":"Elsevier Ltd","title":"Children's engagement with educational iPad apps: Insights from a Spanish classroom","type":"article-journal","volume":"71"},"uris":["http://www.mendeley.com/documents/?uuid=acb69123-fe79-4ce4-9616-f033c6f7fe7c"]}],"mendeley":{"formattedCitation":"(Kucirkova, Messer, Sheehy, & Fernández Panadero, 2014). There is a proliferated tendency in the direction of using mobile devices from the young children, because their touch screens are exactly the proper size for a young child to deal with.

The apps which are constructed to be performed in touch-screen devices, are particularly suitable for young children. What affects young children on considering them appealing, simple and intuitive to use is mainly the playful features of these apps (Dorouka et al., 2020a). Smart mobile devices and apps can promote new and effective teaching practices when encouraging general open-ended exploration and scaffolding experiences. Therefore, the use of touch-screen technologies seems to be necessary, as their integration to early childhood environments may urge young children's multimodal learning experiences in linguistic, spatial, visual, gestural, aural, communicative manners.

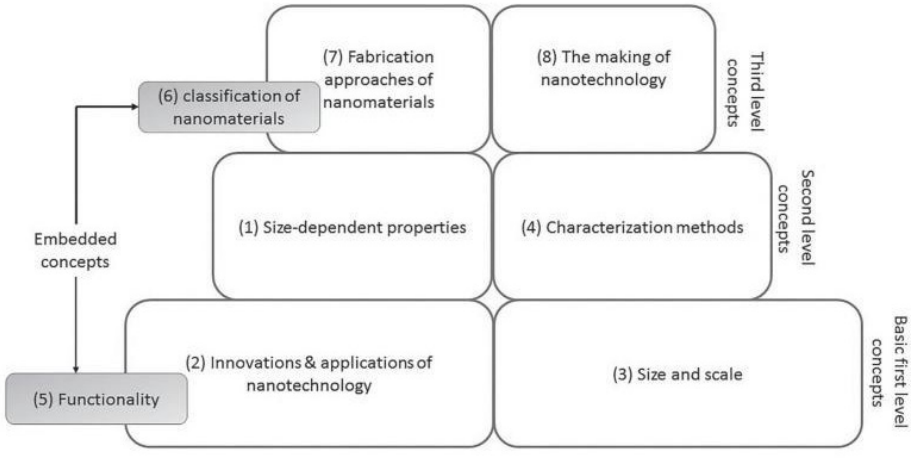
Theoretical Background

We live in a society based not only on technology but science as well. A fundamental part of science concerns elements that cannot be noticed by a naked eye and, undoubtedly, require other ways of teaching in order to enable children to understand them. These distinct aspects of Science are well aligned with the capabilities of smart mobile devices, the key feature of which is their ability to display interactive and three-dimensional simulations (Zydney & Warner, 2016).

Nanoscience and Nanotechnology (NST) seems to be an advantageous context for the children's improvement for knowledge which is related to science education (Blonder & Sakhnini, 2012). NST has been identified as a developing technology of the 21st century, owing to the fact that it combines various scientific subjects, such as Science, Technology, Chemistry, Biology, Art, Mathematics and Engineering and, at the same time (Mandrikas et al., 2019)while RRI issues are discussed through newspaper articles and students' interaction with nanoscientists. Sample: 45 6th Grade primary students from Greece took part in the study during the academic school year 2014–15. Design and methods: The teacher who taught the TLS also assisted in creating its design through collaborating with experts from the field of NST, science education and science communication. Data was collected through students' worksheets and interviews with students. For the data analysis, qualitative methods of content analysis were used. Results: The results showed that primary students can successfully understand NST key concepts like 'size and scale' and 'size-dependent properties'. Regarding the dimensions of RRI, NST appear to be a useful context for enhancing students' discussions more about Ethics, Engagement and Science Education and less about issues related to Gender Equality, Open Access and Governance. Conclusions: This study demonstrates that teaching NST is worthwhile even in primary education. Moreover, NST provide an appropriate context for the introduction of RRI in science education.","author":[{"dropping-particle":"","family":"Mandrikas","given":"Achilleas","non-dropping-particle":"","parse-names":false,"suffix":""},{"dropping-particle":"","family":"Michailidi","given":"Emily","non-dropping-particle":"","parse-

names":false,"suffix":""},{"dropping-particle":"","family":"Stavrou","given":"Dimitris","non-dropping-particle":"","parse-names":false,"suffix":""}],"container-title":"Research in Science and Technological Education","id":"ITEM-1","issue":"00","issued":{"date-parts":[["2019"]]},"page":"1-19","publisher":"Routledge","title":"Teaching nanotechnology in primary education","type":"article-journal","volume":"00"},"uris":["http://www.mendeley.com/documents/?uuid=93860ed4-3f2d-41cb-8020-acf29fddf82e"]}],"mendeley":{"formattedCitation":"(Mandrikas et al., 2019. NST promises to resolve global intertemporal challenges related to the diagnosis and treatment of diseases, the environment, energy, space and more. In this way, it is expected to strongly influence every human's life (Lin et al., 2015).

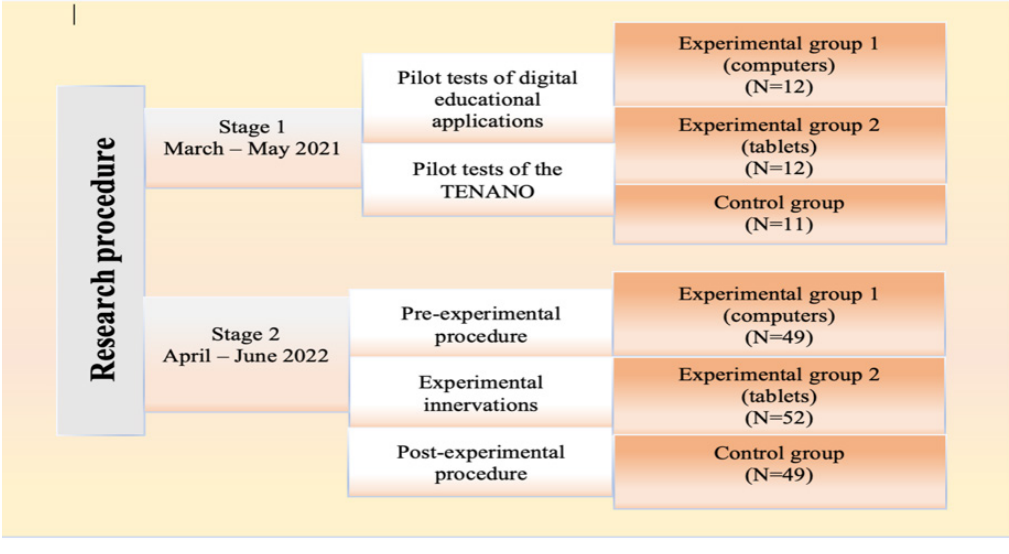
Figure 1. Nano-literacy through the key concepts/ideas of nanotechnology (Adapted from Blonder & Yonai, 2020).



The first step in this direction is the introduction of NST concepts on young children's education through appropriately designed digital tools. Educators have begun identifying Big Ideas (BI) linked to the teaching of nanotechnology (Blonder, & Sakhnini, 2012; Stevens et al., 2009) (see Figure 1). One barrier to children beginning to understand NST is the lack of sensory experience of the non-visible world; because NST focuses on studying the properties of matter and manipulating it at scales ranging from 1-100mn (Xie & Pallant, 2011) and is counterintuitive to common intuition. Considering that there is a strong dependence on the concepts of size and scale on children's sensory experience (Blonder and Saldmini, 2020), a promising way to strengthen the understanding of the continuum in the nature of scale is to visualise objects with significant differences in size (Stavrou et al., 2018). In this way, the distinct aspects of NST are well-aligned with the capabilities of smart mobile devices. Through

visualisation, children can easily perceive the relative difference between different objects in size (Dorouka et al, 2021b). Using technology as a learning tool for the morphology of non-visible objects could be beneficial (Dorouka et al, 2020b). The features that make smart-screen devices, such as tablets, potentially ideal tools, especially for NST teaching, are their characteristics such as the ease of use, the speed (Dorouka et al., 2021a), the flexibility concerning the location where the learning process takes place, the possibility of immediate information and feedback and the possibility of representation of both natural and semantic space or augmenting it with digital information. Although there are mobile applications (apps) that effectively motivate young children's involvement with STEAM-oriented activities, such as Scratch 3, no mobile app was found in the literature that supports properly the young children's learning engagement to the distinct aspects of Science that cannot be noticed by a naked eye, such as NST.

Figure 2. Research design



Method

Given the importance of NST, this study examines whether early-primary school students can benefit from integrating digital technology (computers and mobile devices) to initiate their nano-literacy. Although research has emerged on new technologies such as tablets in early childhood education (Chmiliar, 2017), there needs to be more data on the practical use of mobile devices in learning Science, Technology, Engineering and Mathematics (STEM) to early childhood. Therefore, the present experimental study aims to address the following research hypotheses:

1. Digital technologies are more developmentally appropriate than traditional

ones for introducing nanotechnology elements to the early school-age children to.

2. The performance of the two experimental groups will differ significantly after the intervention, depending on the teaching technique even after controlling the gender of the children, their level of non-verbal cognitive ability and their initial level of nano-literacy.

To investigate the research hypotheses, a teaching intervention was carried out to 150 children of the 2nd grade during one week (see Figure 2). More specifically, three groups were created which approached fundamental elements of NST (see Table 1) - the first experimental group using educational software via computer, the second experimental group using the same educational software running on smart mobile devices (tablets) (see Table 2) and the group with alternative experiential teaching (control group) (see Table 3).

Table 1. The educational intervention for the 3 groups of children

Objectives-Expected learning outcomes	**Activity: Title & short description**	**Content of activities**
Ordering of objects in the macrocosm, microcosm and nanoworld based on criterion 1 **Grouping of** objects in the macrocosm microworld and nanoworld based on criterion 2, 'which objects fit into or are part of other objects.' **Explanation of viral infection**s as a direct process based on a sequence of stages involving macro-, micro- and nanoworld objects (e.g., viruses, cells, humans)	1st: Prometheus and the gift of fire: According to the **Artwork "Days and Theogony" of Hesiod**, out of **Pandora's box** pour out the suffering of people, such as viruses, but also hope! 2nd: **Who is troubling Nana?** Nano world of course: Nana, a little girl, has symptoms of viral infection, such as coughing and fever. 3rd: **Super heroes in medical action**: Nana needs you! A virus has entered her body and is trying to make her sick… Help her!	Children: -express their perceptions of invisible **entities** -interact with the **entities** of the macro-, micro- and nano-world -discover information about the **entities** with which they interact -discover through game/play the criteria for grouping and ordering **entities** in the macro-, micro- and nanoworld -discover through game/play aspects of nature and the role of visualisation in representing invisible **entities** -use the visualised representation of entities in the macroworld, microworld and nanoworld through their correct movement in game/play. -describe invisible **entities** -are encouraged to reflect on their initial perceptions and discuss the visualised **entities** -express their initial perceptions about viral infections -discover information about the behaviour of viruses and the infections they cause in the body through game/play -inquire how viruses, affects the cells as well as the human body. -interact by representing invisible **processes** -present their behaviour in the game in front of the class to explain viral infections -are encouraged to reflect on their initial perceptions -discuss the visualised **entities** in the game/play

Table 2: The presentation of the design of the Digital game "Super heroes in medical action" based on gamification elements

The digital game "Super heroes in medical action"			
Elements of Gamification features			
Motivation theory	**Gamification elements**	Learning strategies	**Teachers' role**
The theory of self-determination Competence - the sense of achievement and success of each individual when interacting with the environment to acquire knowledge and skills. - Positive feedback and well-measured challenges/stages 1 2 3 Autonomy - the individual's willingness to perform a particular task. - Many choices and tasks without controlling or restraining them, no use of rewards for feedback Social relatedness - the feeling connected to significant others - being part of a 2–3-member group to play	Game interface design elements **- points, badges, and leaderboards/ total scores** Game design mechanics **- time limitations and competitive play** Game design principles **-goals, objectives, playstyle** Game models **- the gaming experience**	Problem-based learning Problem-based learning 1.Engagement: The situation/ problem is made known to the students. -Batman: *Nana needs you; a virus has entered her body; help her!* 2.Investigation: An analysis of the situation is conducted, confirming what information they have/need etc. -Joker: *I don't think that's how you drive the virus into the white cells.* 3.Problem Resolution: All information, options and actions are carefully analyzed. -Catwoman: *With your help, the white cells found the viruses in the body and eliminated them!* 4.Reflection: Review and discussion of the learning process followed and how it could be useful in similar situations. -Batman: *Congratulations, you cleared all the viruses from the body, and Nana is fine!*	Supporting Motivational Facilitative

Table 3. Key phases of the theatre play according to Kouretzis (2008)

	Description of phase	Facilitator's encouragement	Characteristics	Materials
A	The 'liberation' phase - the stage of awareness and group formation	-Let's sit in a semicircle, think and express what might be behind the cloth. […] -Time to find out what is behind the cloth by removing it from its place!	physical expression improvisation	colourful cardboard colourful gymnastic hoops
B	The phase of reproduction - a stage that focuses on role-playing and role composition	-Take a hat from the box and look at the message "hidden" to discover your role in the game. • white hats: white cell's role (protect the human's body from enemies such as viruses) • red hats: red cell's role (circulate in human's body and help them breathe) • green hats: virus's role (enter human's body and after hiding inside the red cells to guard against the white cells and reproduce, they travel through the body) -Hats of the same colour form a group.	the exploitation of chance spontaneity	paints pencils rope printed material cardboard boxes rags balls
C	Phase improvisation - Performance phase	-Change groups, and each gets a hat and a wreath of a different colour. -Show the journey of the virus inside Nana's body. -Listen to advice from the superheroes Batman and Catwoman! -Be careful of Joker, Batman and Catwoman's enemy; he is on the side of the viruses that want to make Nana sick! -Who will ultimately win, the white cells or the viruses?		
D	Analysis phase - multi-level processing stage	-Ask for clarification from the others who played, and give your opinion on the roles, the "solution" you chose for the virus or the cells inside Nana's body to win.		

Before and after the end of the teaching interventions, all three groups were tested to see whether there were any differences in the level of children's nano-literacy. In particular, each group went through three phases. The first and the third phase

involved individual semi-structured interviews with the same questions for each child, whereas the second phase referred to the teaching NST elements.

Table 4. Examples of "good answers" of the children on TENANO after the intervention

TENANO Questions		Examples of Children's Responses
1	What is the **minor thing** you can think of?	Virus, coronavirus
2	**Order** what the pictures show from the biggest to the smallest.	man<ball<ant<white cell<red cell<virus
3	**Put** the pictures of things that look alike in how small or big they are **into groups.** -If one is a piece of another, put those 2 in 2 different groups.	{man, ball, ant} {white cell, red cell}, {virus}
4	Two siblings **had a virus with a cough and a fever** and did not go to school for a week. When they returned to school, their classmates asked them **why** they did not come. What would you answer in their place? Describe the virus's journey through the body using the pictures from the previous question that appear as a virus, cells or the human body.	*I got sick, and I couldn't come to school; the virus got inside me, and it went on its journey, and it was hiding in the red cells and I made other friends like it, but the white cells saw it and fought it. I got better.*

As part of the interview, the children were given an appropriately designed test (**TENANO**) that measured their knowledge of comparing, ordering and grouping macro-world, micro-world and nano-world elements (see Table 4) and that was created after a thorough review of the literature (Delgato et al., 2015; Magana et al., 2012; Stevens et al., 2009).

Results

A x^2 test was conducted to test the gender equivalence between the three groups. No difference was found at a statistically significant level between the number of boys and girls, $x2(2) = 0.30$, $p > 0.05$. In order to discover the differences between the groups in terms of level non-verbal cognitive ability and their initial level of nano-literacy, a one-way analysis of variance (ANOVA) was performed. Results showed that the experimental and control groups did not differ in terms of nonverbal cognitive ability level ($F(2, 149) = 2.62$, $p > .05$), or TENANO scores before the intervention ($F(2, 149) = 1.62$, $p > .05$), at a statistically significant level.

Table 5. Means (M), standard deviations (sd) of the TENANO test after the teaching intervention by group and results of the analysis ANOVA

	Nano-literacy post-test items – TENANO				
	Minor thing	**Ordering**	**Grouping**	**Explanation of infection**	**TOTAL SCORE**
	M (SD)	**M (SD)**	**M (SD)**	**M (SD)**	**M (SD)**
Experimental (PC)	1.96(.20)	2.67(.62)	1.24 (.59)	1.78 (.55)	7.65 (1.18)
Experimental 2 (tablets)	1.87(.39)	2.83(.51)	1.90 (.35)	1.88 (.42)	8.48(0.89)
Control group	.94(.94)	.45(1.04)	.29 (.45)	.51 (.82)	2.18(1.92)
ANOVA	F(2, 149) =43.71, $p < .05$	F(2, 149) =153.08 $p < .05$	F(2, 149) =145.72 $p < .05$	F(2, 149) =75.90, $p < .05$	F(2, 149) = 299.74, $p < .05$

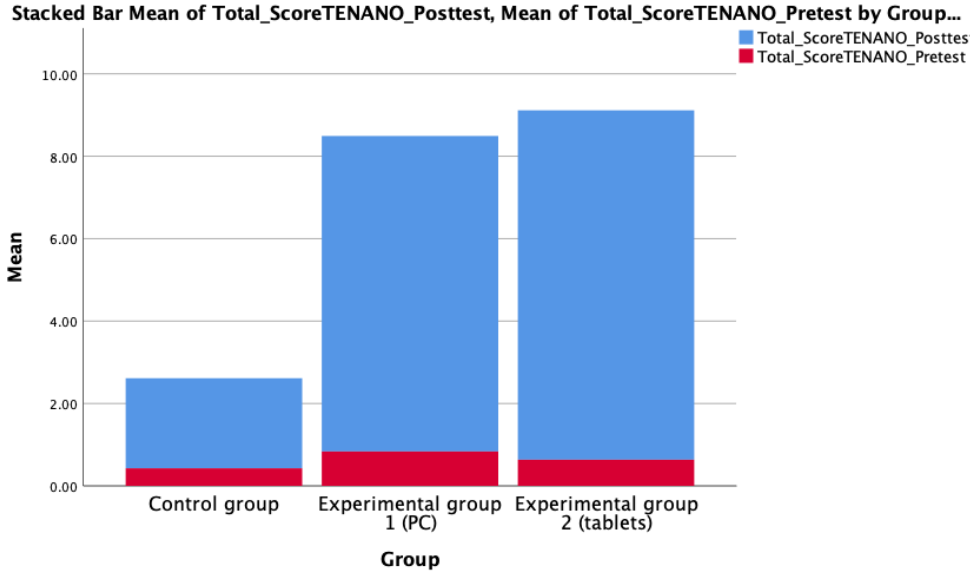
Moreover, one-way analysis of variance (ANOVA) indicated that the main effect of the independent variable (group type) on children's performance on TENANO after the teaching intervention was found to be statistically significant, ($F(2, 149) = 299.74$, $p < 0.001$) the group with alternative experiential teaching (control group) (see Table 5). Both experimental and control groups increased the children's performance on TENANO after the completion of the experimental intervention. The learning achievements of the students in the second experimental group (tablets) were significantly better than the ones of the students in the first experimental group (computers) and the students in the control group.

However, the results of the analysis of the items of TENANO showed that the performance of children in the experimental group 2 (tablets) did not differ statistically significantly compared to the performance of children in experimental group 1 with computers in some cases. The performance of the children in the experimental group with computers on the TENANO assignment task involving the smallest thing, ordering and explanation of the virus infection was not statistically significantly higher than the performance of the children in the experimental group 2 (tablets) ($p>.05$). In contrast, statistically significant differences were found between the experimental group 1 (computers) and the experimental group 2 (tablets) ($p<.05$) in their grouping and overall performance (see Table 6) on each item of the TENANO, with the tablet group being superior. Statistically significant differences were found between the experimental group 1 (computers) and the control group ($p<.05$) in their performance on each item of TENANO, with experimental group 1 significantly outperforming the control group.

Analyzing further the process data, it must be noticed that the learning outcomes

were positive for all the research groups (see Figure 3). Although, in several cases, the students in the computer group had difficulties in using the keyboard and mouse, they reached the completion of the digital activities as well as the children in the tablet group. The students in the control group enjoyed the theatre play activity and their enthusiasm for participation was high (see Figure 4). Therefore, it is concluded that the learning approach using digital technologies to introduce NST elements significantly improved the learning outcomes of young children (Dorouka & Kalogiannakis, 2023).

Figure 3. TENANO scores before and after the teaching intervention for each group



The achievability of young children's nano-literacy development is confirmed by other research. While Voo & Lajium's (2022) study agrees that students' understanding of the concept of size and scale involving reasoning ability is a high cognitive ability, Mandrikas et al. (2020), as well as Peikos et al. (2022), have demonstrated that students as early as elementary school can reach this cognitive level when teachers utilise readily available instructional materials to help students understand abstract concepts.

Table 6. Post Hoc Tests for Experimental 1 PC (EG1_PC), Experimental 2 Tablets (EG2_TABLETS) and Control Group (CG) after the intervention

Total Score of TENANO				
EG1_PC	EG2_TABLETS	-.82771*	.27743	.010
	CG	5.46939*	.28152	.000
EG2_TABLETS	EG1_PC	.82771*	.27743	.010
	CG	6.29710*	.27743	.000
CG	EG1_PC	-5.46939*	.28152	.000
	EG2_TABLETS	-6.29710*	.27743	.000

*. The mean difference is significant at the 0.05 level.

Another study by Lin et al. (2015) revealed that primary school students who participated in a nanoscience camp to promote science education, students had achieved significant improvement in their understanding of NST topics after participating in interactive, hands-on activities.

In addition, we attempted to investigate possible gender differences in children's performance in nano-literacy. We did not find that gender played a significant role, as we did not find statistically significant differences in learning performance between boys and girls. The gender-related results of the study are consistent with existing literature (Hyde & Mertz, 2009; Yang & Bers, 2023; Gunbatar et al., 2018), as they suggest that in young children, gender differences in science performance are either non-existent or very small (Goodwin, 2012).

Figure 4. Illustrations from the activities of children



Discussion & Conclusion

This study is the effects of different forms of digital technology for the understanding of NST elements. From the results, the first significant finding was that the performance of the tablet group was higher than the one of the computer group and the control group. This confirmed the findings of previous research which compared the potential of digital technologies in terms of their

effectiveness in early childhood (Papadakis et al., 2016). In addition, these findings suggest that early childhood children can learn through interactive play experiences from touchscreen mobile devices. There is, also, evidence to support the use of mobile apps in learning programmes and the claim that, if there are used properly, they enhance learning in STEM, reading and writing skills, and young children's motivation to engage in the learning process (Herodotou, 2018).

However, as can be seen from the results (see Table 6) of the lower performance of experimental group 1 compared to experimental group 2 on the grouping task on the one hand and the overall TENANO score on the other hand, the use of computers had a more negligible effect than the use of tablets. Tablets were more effective than computers due to the former's intuitive touch-based interface (Neumann, 2018) for discovering the criteria for grouping entities in the macro-, micro- and nano-world through the game, capturing the importance of comparing the two digital technologies. That is, the more passive role of children when engaging with computers compared to engaging with tablets probably did not sufficiently encourage children to actively engage with the information provided by the activities carried out by the digital applications to correctly use the visualized representation of macro-, micro- and nano-world entities through their correct movement in the game. The present study is consistent with other studies, which show that the interactive environment created in an elementary school using tablets is crucial to maintain children's interest in digital activities (Liu, 2013; Risconscente, 2012). Moreover, it encourages them to engage more effectively and closely with digital activities for learning scientific or interdisciplinary concepts.

Limitations - Future Research

Among the limitations of this study was the short duration of the teaching intervention. Although it was sufficient to experimentally test the effect of different teaching techniques, more was needed to meet the needs of young children in best developing their level of nano-literacy. Therefore, it is necessary to implement a more detailed teaching intervention compatible with the NST field and of sufficiently long duration to extensively investigate the effect of different teaching approaches on further developing the ability to understand fundamental NST concepts.

A complete digital nano-literacy tool can be for every young member of the modern multicultural society, the educational and socioeconomic "cure" of our time. Actions in health and education can alleviate the current stagnation of the economy. In this struggle, each member of the global society can stand as essential allies, so it is imperative that they can understand the basic NST concepts. An expected outcome of the situation is an increased interest in STEM professions,

as nano-literate members of society will have improved employability, which aligns with the philosophy of distance education imposed by the recent health crisis and the strong wave of migration.

Ethics Statement

Before the start of the first phase of the research, the research protocol was reviewed and approved by the Ethics Committee of the Department of Preschool Education of the University of Crete (no 135/24-03-2020 & 1016/22-11-2021), the Ethics and Research Ethics Committee of the University of Crete (REC-UOC) (no. 35/24.02.2022, https://www.ehde.uoc.gr/index.php/en), as well as by the Hellenic Institute of Educational Policy (IEP) (no. 3578/14-03-2022).

Disclosure Statement

No potential conflict of interest was reported by the authors.

Acknowledgement

The research work was supported by the Hellenic Foundation for Research and Innovation (HFRI) under the 3rd Call for HFRI PhD. Fellowships (Fellowship Number: 5503).

References

Blonder, R., & Sakhnini, S. (2012). Teaching two basic nanotechnology concepts in secondary school by using a variety of teaching methods. *Chemistry Education Research and Practice*, *13*(4), 500-516. https://doi.org/10.1039/C2RP20026K

Blonder, R., & Yonai, E. (2020). Exposing school students to nanoscience: A review of published programs. In K. D. Sattler (Ed.), *21st Century Nanoscience – A Handbook: Public Policy, Education, and Global Trends* (Vol. 10). Boca Raton: Taylor & Francis (CRC Press). https://doi.org/10.1201/9780429351631

Chmiliar, L. (2017). Improving learning outcomes: the iPad and preschool children with disabilities. *Frontiers in Psychology*, pp. *8*, 660. https://doi.org/10.3389/fpsyg.2017.00660

Delgado, C., Stevens, S.Y., Shin, N. & Krajcik, J. (2015). A middle school instructional unit for size and scale contextualized in nanotechnology, *Nanotechnology Reviews*, *4*(1), 51–69. https://doi.org/10.1515/ntrev-2014-0023

Dorouka, P., & Kalogiannakis, M. (2023). Teaching Nanotechnology Concepts in Early-Primary Education: An Experimental Study Using Digital Games. *International Journal of Science Education*. https://doi.org/10.1080/09500693.2023.2286299

Dorouka, P., Papadakis, St., & Kalogiannakis, M. (2020a). Tablets & apps for promoting Robotics, Mathematics, STEM Education and Literacy in Early Childhood Education, *International Journal of Mobile Learning and Organisation*

(IJMLO), *14*(2), 255-274. https://doi.10.1504/IJMLO.2020.10026334.

Dorouka, P., Papadakis, St., & Kalogiannakis, M. (2020b). The influence of digital technology on young children's "nano-literacy". In K. Plakitsi, E. Kolokouri & A.-C. Kornelaki (Eds.), *ISCAR (International Society of Cultural-historical Activity Research) Regional Conference 'Crisis in contexts', e-proceedings*, 308-320, University of Ioannina, 19-24 March 2019.

Dorouka, P., Papadakis, St., & Kalogiannakis, M. (2021a). Nanotechnology and Mobile Learning: Perspectives and Opportunities in Young Children's Education, *Int. J. Technology Enhanced Learning, 13*(3), 237-252. https://doi.10.1504/IJTEL.2021.115975

Dorouka, P., Papadakis, St., & Kalogiannakis, M. (2021b). The contribution of the health crisis to young children's nano-literacy through STEAM education, *Hellenic Journal of STEM Education*, *2*(1), 1-7. https://doi.org/10.51724/hjstemed.v2i1.18

Goodwin, K. (2012). Use of Tablet Technology in the Classroom. *Education and Communities*, 1–96.

Herodotou, C. (2018) Young children and tablets: A systematic review of effects on learning and development. *Journal of Computer Assisted Learning. 34*(1), 1–9. https://doi.10.1111/jcal.12220

Kouretzis, L. (2008). *The play and its dimensions*. Athens: Taxideftis. [in Greek]

Kucirkova, N., Messer, D., Sheehy, K., & Fernández Panadero, C. (2014). Children's engagement with educational iPad apps: Insights from a Spanish classroom. *Computers and Education*, *71*, 175–184. doi: 10.1016/j.compedu.2013.10.003

Lin, S. Y., Wu, M. T., Cho, Y. I., & Chen, H. H. (2015). The effectiveness of a popular science promotion program on nanotechnology for elementary school students in I-Lan City. *Research in Science and Technological Education*, *33*(1), 22–37. https://doi.10.1080/02635143.2014.971733

Magana, A. J., Brophy, S. P., & Bryan, L. A. (2012). An Integrated Knowledge Framework to Characterize and Scaffold Size and Scale Cognition (FS2C). *International Journal of Science Education*, *34*(14), 2181–2203. https://doi.10.1080/09500693.2012.715316

Mandrikas, A., Michailidi, E., & Stavrou, D. (2020). Teaching nanotechnology in primary education. *Research in Science & Technological Education*, *38*(4), 377–395. https://doi.org/10.1080/02635143.2019.1631783

Mandrikas, A., Michailidi, E., & Stavrou, D. (2019). Teaching nanotechnology in primary education. *Research in Science and Technological Education*, *00*(00), 1–19. https://doi.10.1080/02635143.2019.1631783

Papadakis, S., Kalogiannakis, M., & Zaranis, N. (2016) Comparing tablets and PCs in teaching mathematics: An attempt to improve mathematics competence in early childhood education, *Preschool and Primary Education*, *4*(2), 241–253. https://doi.10.12681/ppej.8779

Peikos, G., Spyrtou, A., Pnevmatikos, D., & Papadopoulou, P. (2022). A teaching

learning sequence on nanoscience and nanotechnology content at primary school level: evaluation of students' learning, *International Journal of Science Education*, *44*(12), 1932-1957. https://doi.org/10.1080/09500693.2022.2105976

Stavrou, D., Michailidi, E., & Sgouros, G. (2018). Development and dissemination of a teaching learning sequence on nanoscience and nanotechnology in a context of communities of learners. *Chemistry Education Research and Practice*, *19*(4), 1065-1080. https://doi.org/10.1039/C8RP00088C

Stevens, S. Y., Sutherland, L. M., & Krajcik, J. S. (2009). *The big ideas of nanoscale science and engineering. Guidebook for Secondary Teachers.* Arlington-VA: NSTA Press.

Voo, L. H., & Lajium, D. A. (2022). Scoping Review: Appropriate Big Ideas of Nano Science and Nanotechnology to Teach in Chemistry for Secondary School. *Malaysian Journal of Social Sciences and Humanities (MJSSH)*, *7*(12), e002016-e002016. https://doi.org/10.47405/mjssh.v7i12.2016

Xie, C., & Pallant, A. (2011). The Molecular Workbench Software: An Innovative Dynamic Modeling Tool for Nanoscience Education. *Springer Netherlands EBooks*, 121–139. https://doi.org/10.1007/978-94-007-0449-7_6

Zhou, S. (2022). Effect of Mobile Learning on the Optimisation of Preschool Education Teaching Mode under the Epidemic, *Wireless Communications and Mobile Computing*, *2022*. https://doi.org/10.1155/2022/2194373

Zydney, J. M., & Warner, Z. (2016). Mobile apps for science learning: Review of research. *Computers and Education*, *94*, 1–17. https://doi.org/10.1016/j.compedu.2015.11.001.

Introducing the Control of Variables Strategy to Early Childhood Student Teachers With the Assistance of an Augmented Pedagogical Agent

Angelos Sofianidis[1], Silvestra Sakellariou[2] Anastasios Zoupidis[3], Euripides Harzikraniotis[2]

[1]University of Western Macedonia, Greece

[2] Aristotle University of Thessaloniki, Greece

[3] Democritus University of Thrace, Greece

During the last decade, Augmented Reality (AR) has contributed to teaching challenging topics of Science Education in several ways. Our paper focuses on introducing the Control of Variables Strategy (CVS) -an essential aspect of inquiry and a challenging topic for students of all ages- using an AR-enhanced approach. According to the existing literature, combining explicit instruction with students' experimentation is one of the most effective approaches in teaching and learning the CVS method. Based on these findings, an Augmented Pedagogical Agent called Nefeli was designed to assist early childhood student teachers in implementing CVS when facing problems with two independent variables in inquiry-based activities. The aim of the present study was to explore students' perceptions regarding their experience with the Nefeli application and its characteristics and assess the learning outcomes of the intervention designed. The results indicate that most students found Nefeli very interesting and useful in understanding how to apply CVS principles. The findings demonstrate the positive impact of an augmented pedagogical agent on enhancing students' learning experiences and outcomes in CVS understanding, while also contributing to the existing research on the relationship between CVS-related skill development and the content, as well as to the discussion about the most effective approach on teaching and learning CVS.

Keywords: Augmented reality, control of variables strategy, early childhood student teachers.

Introduction

Augmented reality (AR) technologies in education have attracted increasing attention from academics and teachers over the past ten years. The technology of AR allows for the overlay or projection of visual items over the physical world (Akçayır & Akçayır, 2017). The literature emphasises the benefits of implementing AR in educational contexts, which, according to empirical studies,

improve teaching and learning outcomes (Chiang et al., 2014; Efstathiou et al., 2018; Garzón, 2021) and enhance students' self-directed learning (Wen et al., 2023) in inquiry-based learning environments. Also, Sofianidis (2022) reports that using AR helped students learn more easily and created a sense of personal communication that was found to be very engaging for students. Based on these findings, AR seems to be an appropriate tool to foster students' learning in challenging subjects or practices, such as the Control of Variables Strategy.

The Control of Variables Strategy (CVS) is an essential aspect of inquiry (Pedaste et al., 2015) whereby students are expected to investigate causal relationships between variables. Research shows that students of all ages, ranging from Primary to Tertiary education, have difficulty implementing CVS and follow unorthodox practices even after specifically designed interventions (Boudreaux et al., 2008; Zoupidis et al., 2017). This difficulty is attributed to the fact that a combination of skills is required, such as logical thinking skills (Chen & Klahr, 1999) and scientific argumentation (Erlina et al., 2018). In addition, students' scientific content knowledge and their general cognitive abilities are essential for the proper implementation of the method (Schwichow et al., 2020). On the other hand, several studies support the idea that combining explicit instruction with implicit learning through experimentation is one of the most effective approaches in teaching and learning the CVS method (Lorch et al., 2010; Zoupidis et al., 2017).

Based on the above findings, we designed a pedagogical agent in an AR environment that offers students explicit instruction on the CVS principles while they try to solve problems through a guided inquiry-based approach. The Augmented Pedagogical Agent is called "Nefeli", and it is an AR-enhanced agent designed to advise students when they engage in questions about the CVS method during inquiry-based activities. The purpose of the current study was to evaluate learning results and investigate students' opinions of their experiences by implementing an AR-enhanced approach with a group of early childhood student teachers.

Methodology

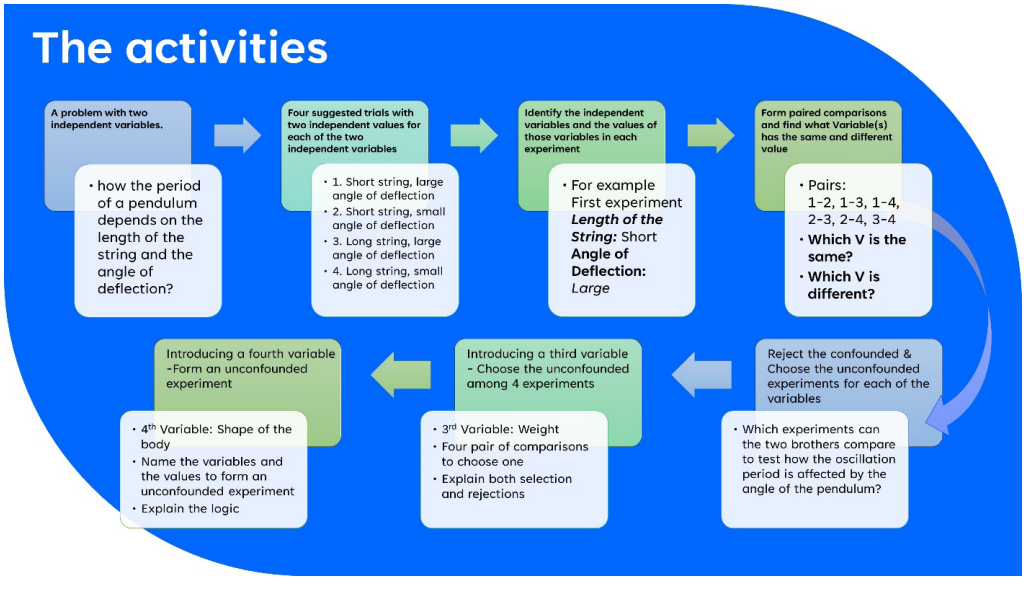
Participants

A total of 65 student teachers participated in the two phases of the current study. Specifically, 37 student teachers were in Phase A (pilot implementation) and 28 in Phase B (main implementation of the study). All students of both phases completed the Nefeli-Questionnaire regarding their experience of using Nefeli and its characteristics. In the case of phase B, only 24 out of 28 student teachers completed all the CVS-Questionnaire assessing their CVS-related skills (pre-, mid-, post-), so only their data were analysed.

Design and Development

Students working in pairs face two inquiry-based problems with Nefeli's assistance in two different lessons. The inquiry-based activities are designed to have the same structure (Figure 1). First, students are introduced to a problem with two independent variables. Each of the two independent variables has two independent values, and four proposed trials are provided. Next, students are asked to identify each trial's independent variables and their values. They are then required to create paired comparisons, i.e., experiments, and think about their validity, pinpointing valid and invalid experiments.

Figure 1. A brief description of the activities' structure



Nefeli was developed on the GoMeta platform and is based on a teacher figure. Nefeli greets them and asks as to which question they require assistance with first. Throughout all activities, Nefeli provides students with non-content-specific advice to help them answer the questions. Nefeli's advice is based on the principles of CVS (e.g., "In a valid comparison, only the focal independent variable can have different values, and all the others have the same value in each trial").

During the activities, the students are prompted to use Nefeli's advice in all steps. Then, with the guidance of the instructor, the whole class uses an interactive online quiz to reflect on the ways in which they managed the variables during the activities in order to highlight the CVS principles in a more explicit way.

Then, students participated in a lesson dedicated to reflecting on the activities.

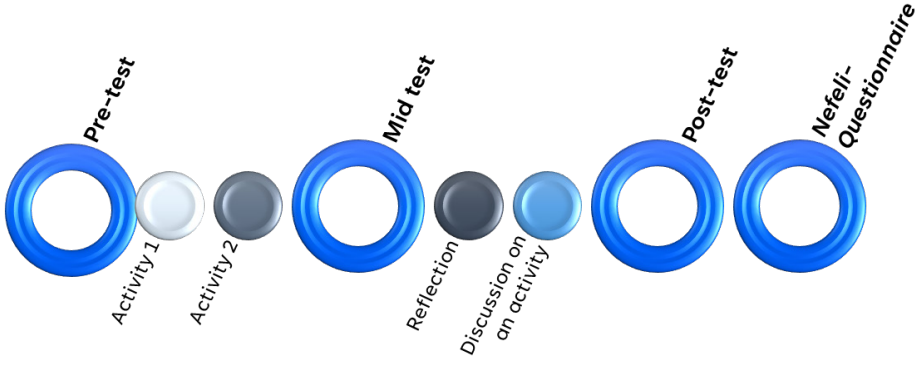
The students, guided by the instructor, explicitly state the CVS principles and discuss one of the abovementioned activities.

Nefeli was developed in two phases. Phase A was a pilot implementation aiming to explore students' perceptions and track down possible adjustments to the application. Phase B was the main implementation of the current study, aiming to assess students' learning outcomes on CVS-related skills and explore their perceptions of their experience with the application and Nefeli's characteristics.

Data collection

The data concerning students' perceptions of their experience were collected using an anonymous Nefeli-Questionnaire composed of 4-Likert scale questions after the end of all activities. The data was collected after piloting (Phase A) and the main implementation (Phase B) of the current study.

Figure 2. The sequence of the activities during the main implementation, including the data collection



Students who participated in the main implementation (Phase B) were also asked to individually complete a CVS-Questionnaire (Zoupidis et al., 2017; Zoupidis et al.,2021) without Nefeli's assistance (Figure 2) aiming to assess their following CVS-related skills:

a) Designing an unconfounded experiment

b) Identifying an unconfounded experiment among four experiments

c) Interpreting an unconfounded experiment, and

d) Understanding the vagueness of a confounded experiment.

Students completed the CVS-Questionnaire for the first time (pre-test) before the use of Nefeli, the second one (mid-test) after the use of Nefeli and the third one

(post-test) after the reflection and the explicit statement of the CVS principles. The CVS-Questionnaire includes questions to assess the abovementioned skills in two physics topics (sinking and floating and magnets), for example, "*Imagine that we put an object into water. A fellow student claims that "the weight of the object affects its floating or sinking, i.e. that heavier objects sink and lighter objects float". We conducted the following four experiments. Which one or which of these experiments can you use to conclude whether your fellow student is right? Why? Why did you reject the other experiments?"*

Data analysis

The data from the Nefeli-Questionnaire regarding students' perceptions of their experience with the Nefeli application and its characteristics were analysed using descriptive and non-parametric statistics (Mann-Whitney U test) to compare the results between the two phases of the study (Phase A and B). Student responses on the CVS-Questionnaire were first qualitatively analysed using a rubric by two independent researchers. Then the results were analysed statistically using descriptive and non-parametric statistics (Wilcoxon test) to explore statically significant differences among students' learning outcomes measured in the pre-, mid- and post-test.

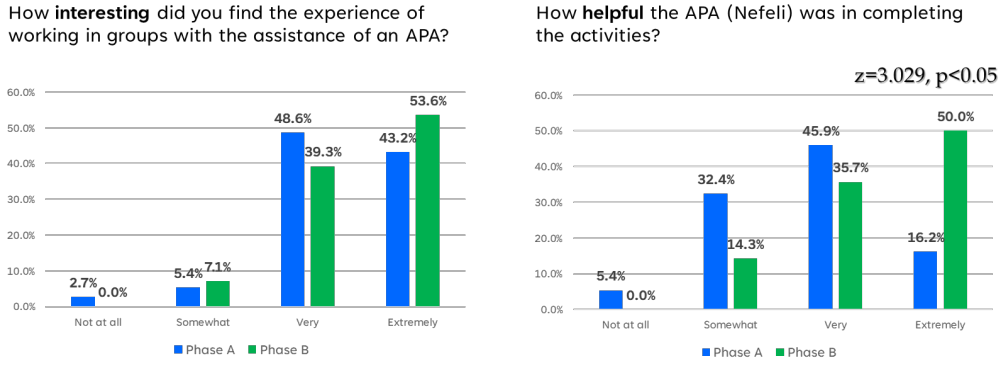
Results

Hereafter, the results concerning students' perceptions of their experience with the Nefeli application and its characteristics are presented, followed by the results regarding students' learning outcomes on CVS-related skills.

Students' perceptions of their experience with Nefeli application and its characteristics

Students who engaged with the second version of the program during the primary implementation phase (Phase B) reported more favourable outcomes across all survey questions. As depicted in Figure 3, participants from both phases expressed that their experience with Nefeli was engaging and that Nefeli was instrumental in facilitating the completion of designated activities. Notably, the difference between the two phases on the enhancement in outcomes related to Nefeli's helpfulness was determined to be statistically significant ($z=3.029$, $p<0.05$).

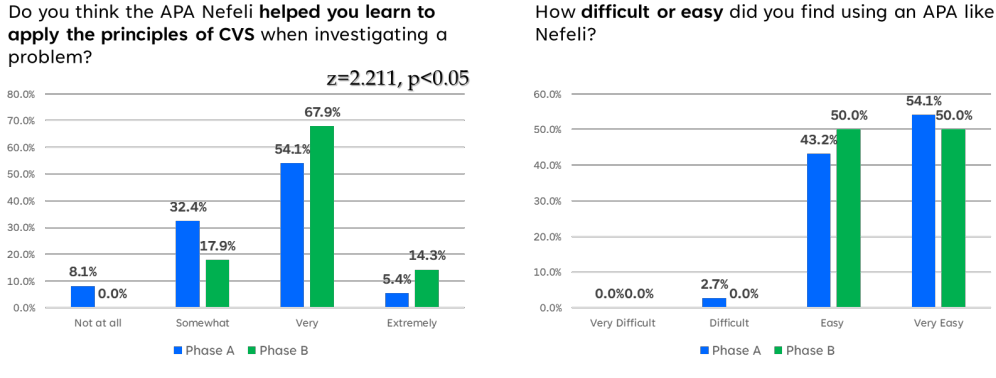
Figure 3. Students' perceptions concerning interest and helpfulness of Nefeli's assistance in the activities



Similarly, this applies to students' perceptions of how the application aided them in learning to apply the principles of CVS, as illustrated in Figure 4. In this respect, the difference between the two phases was also determined to be statistically significant. Moreover, students from both phases reported ease of use with Nefeli.

The characteristics of Nefeli, particularly those related to its medium – including multimodality, augmented reality, and a game-like environment (as shown in Figure 5) – appear to have positively influenced the students' experiences, according to their perceptions.

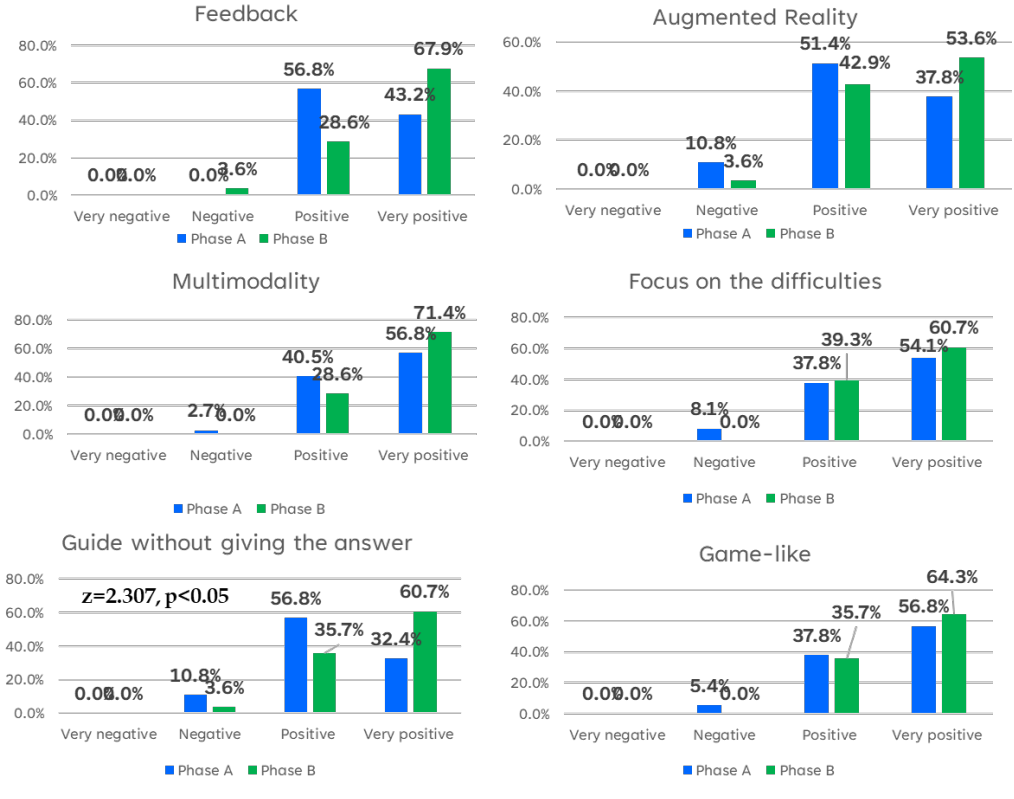
Figure 4. Students' perceptions concerning the helpfulness of Nefeli in learning CVS principles and the difficulty of using Nefeli



This observation is consistent with design elements such as providing feedback, addressing difficulties and providing guidance without revealing the answers. Notably, the results pertaining to the latter showed a significant improvement in

the second phase (z=2.307, p<0.05).

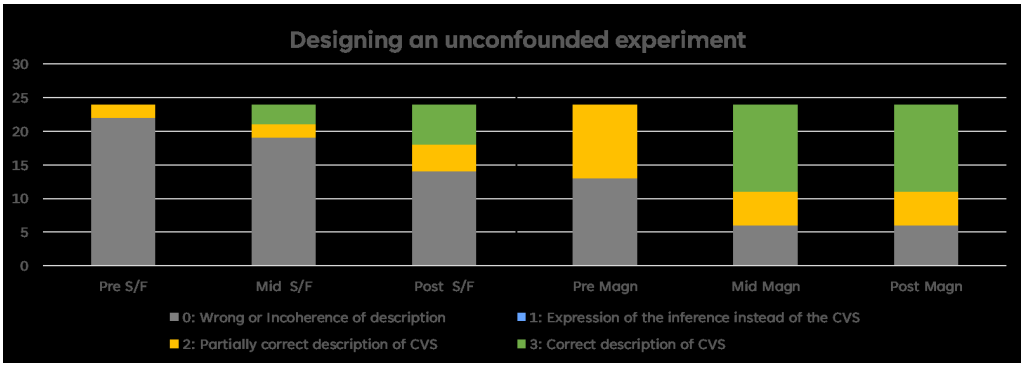
Figure 5. Students' perceptions concerning the main characteristics related to the medium and the design of Nefeli



CVS learning outcomes

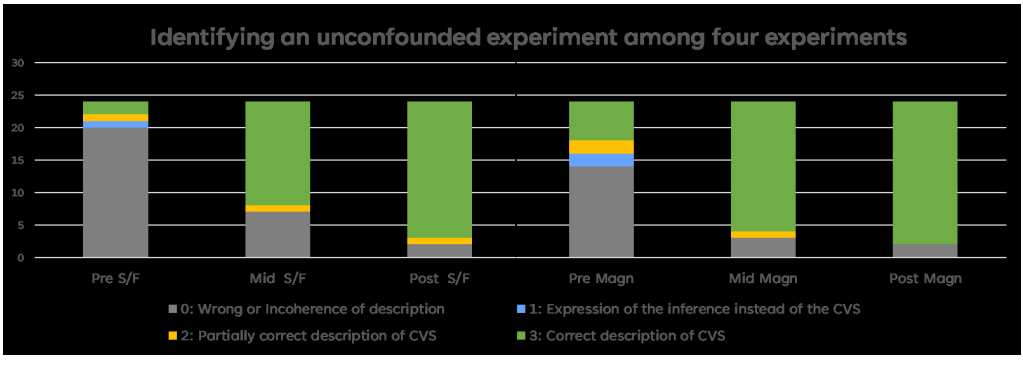
Concerning the first skill, designing an unconfounded experiment (Figure 6), a statistically significant improvement was observed following the intervention across both topics (pre-post: Zf/s:2.724; Zmagn:2.719, p<0.05). Specifically for the topic of Magnetism, the enhancement was statistically significant mid-intervention (pre-mid: Zmagn:2.891, p<0.05), wherein students engaged autonomously with Nefeli. Statistical significance was also found in the differences between the two topics at the pre-, mid-, and post-test stages (Zpre:2.714; Zmid:3.112, Zpost:2.336, p<0.05).

Figure 6. Students' responses concerning designing an unconfounded experiment in pre-, mid-, and post-test



As depicted in Figure 7, a substantial majority of students demonstrated the ability to apply CVS in the case of the second skill, that is, identifying an unconfounded experiment among four experiments (Figure 7), in both topics post-intervention (pre-post: Zf/s:4.233; Zmagn:3.699, $p<0.001$). Additionally, a statistically significant improvement was observed in both topics mid-intervention (pre-mid: Zf/s:3.624; Zmagn:3.573, $p<0.001$). In that case, statistically significant differences between the two topics were only evident pre-intervention (Zpre:2.549 $p<0.05$).

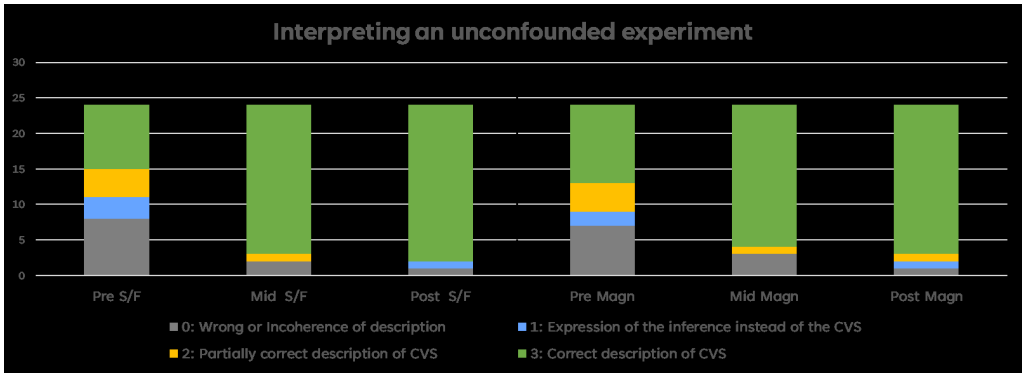
Figure 7. Students' responses concerning identifying an unconfounded experiment among four experiments in pre-, mid-, and post-test



The same applies to the third skill, interpreting an unconfounded experiment (Figure 8). A vast majority of the students successfully applied CVS post-intervention (pre-post: Zf/s:3.345, $p<0.001$; Zmagn:2.987, $p<0.05$), with statistically significant improvements mid-intervention in both topics (pre-mid: Zf/s:3.219, $p<0.001$; Zmagn:2.709, $p<0.05$). In this instance, no statistically

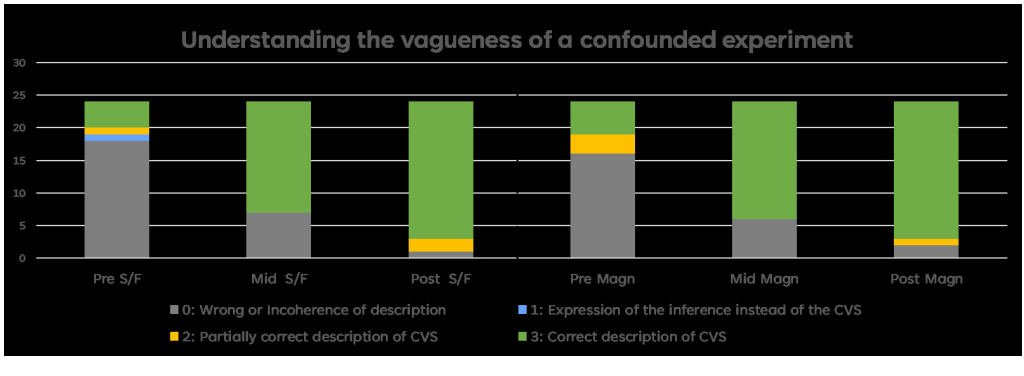
significant differences were detected between the two topics.

Figure 8. Students' responses concerning interpreting an unconfounded experiment in pre-, mid-, and post-test



Regarding the fourth skill, understanding the vagueness inherent in a confounded experiment (Figure 9), the vast majority of students correctly applied CVS (Figure 9) in both topics post-intervention (pre-post: Zf/s:4.062; Zmagn:3.824, $p<0.05$). Moreover, a statistically significant enhancement was noted in both topics mid-intervention (pre-mid: Zf/s:3.136; Zmagn:3.052, $p<0.05$). There were no statistically significant differences observed between the two topics.

Figure 9. Students' responses concerning understanding the vagueness of a confounded experiment in pre-, mid-, and post-test



Discussion

The aim of the present study was to explore students' perceptions regarding their experience with the Nefeli application and its characteristics and assess the learning outcomes of the intervention designed. The final version of Nefeli was improved, compared to the previous one, based on students' responses. In

line with the literature, the augmented pedagogical agent Nefeli seems to be welcomed by the students and helps them complete the activities (Garzón, 2021) and learn how to apply CVS.

After the intervention, the vast majority of the students were able to apply CVS in order to identify and interpret an unconfounded experiment and understand the vagueness of a confounded one. It is important to highlight that these skills were significantly improved even before the last part of the intervention, when students had only worked autonomously with the assistance of Nefeli. The intervention also seems to improve students' ability to design unconfounded experiments. According to the results, another interesting finding is that the development of CVS skills differs depending on the topic.

The positive results concerning the CVS learning outcomes of our intervention seem to support explicit instruction as an effective way of teaching and learning the CVS (Lorch et al., 2010; Zoupidis et al., 2017) but, at the same time, raise questions concerning the need to combine it with experimentation, since students seem to enhance most of their skills without being involved in experiments.

Conclusion

The results of the current study highlight the positive role of the augmented pedagogical agent on students' learning experience and the effectiveness of the proposed series of activities on students' learning outcomes concerning CVS. Furthermore, the findings contribute to the research in progress on two notable teaching and learning CVS aspects. The first aspect pertains to the relation between the development of CVS skills and the topic, i.e., the content, of the teaching and learning activities that are meant to boost this development. The second aspect involves examining whether students require direct involvement in hands-on experiments to develop CVS skills or if a series of preparatory activities similar to those employed in this study could effectively precede their engagement in experiments in which they will be able to apply CVS.

References

Akçayır, M., & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review*, *20*, 1-11

Boudreaux, A., Shaffer, P., Heron, P., & McDermott, L. (2008). Student understanding of control of variables: Deciding whether or not a variable influences the behavior of a system. *American Journal of Physics, 76*(2), 163-170.

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child development*, *70*(5), 1098-1120.

Chiang, T. H., Yang, S. J., & Hwang, G. J. (2014). An augmented reality-based mobile learning system to improve students' learning achievements and motivations

in natural science inquiry activities. *Journal of Educational Technology & Society*, *17*(4), 352-365.

Efstathiou, I., Kyza, E. A., & Georgiou, Y. (2018). An inquiry-based augmented reality mobile learning approach to fostering primary school students' historical reasoning in non-formal settings. *Interactive Learning Environments*, *26*(1), 22-41.

Erlina, N., Susantini, E., Wicaksono, I., & Pandiangan, P. (2018). The Effectiveness of Evidence-Based Reasoning in Inquiry-Based Physics Teaching to Increase Students' Scientific Reasoning. *Journal of Baltic Science Education*, *17*(6), 972-985.

Garzón, J. (2021). An Overview of Twenty-Five Years of Augmented Reality in Education. *Multimodal Technologies and Interaction*, *5*(7), 37.

Lorch, R. F., Lorch, E. P., Ca lderhe ad, W. J., Dunlap, E. E., Hodell, E. C., & Freer, B. D. (2010). Learning the control of variables strategy in higher and lower achieving classrooms: Contributions of explicit instruction and experimentation. *Journal of Educational Psychology, 102*, 90-101.

Pedaste, M., Mäeots, M., Siiman, L. A., De Jong, T., Van Riesen, S. A., Kamp, E. T., Constantinos C. Manoli, Zacharias C. Zacharia... & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational research review*, *14*, 47-61.

Schwichow, M., Osterhaus, C., & Edelsbrunner, P. A. (2020). The relation between the control-of-variables strategy and content knowledge in physics in secondary school. *Contemporary Educational Psychology, 63*, 101923.

Sofianidis, A. (2022). Why Do Students Prefer Augmented Reality: A Mixed-Method Study on Preschool Teacher Students' Perceptions on Self-Assessment AR Quizzes in Science Education. *Education Sciences*, *12*(5), 329.

Wen, Y., Wu, L., He, S., Ng, N. H. E., Teo, B. C., Looi, C. K., & Cai, Y. (2023). Integrating augmented reality into inquiry-based learning approach in primary science classrooms. *Educational technology research and development*, 1-21.

Zoupidis, A., Strangas, A., Kariotoglou, P. (2017). The influence of explicit instruction in understanding the control of variables strategy: the case of preschool students. In B.C. Dimov (Ed.), *Education Across Borders: «Education and research across time and space»* (pp. 412 – 418). Bitola, FYROM: ISBN ISBN 978-9989-100-51-2.

Zoupidis, A., Tselfes, V., and Kariotoglou, P. (2021). Pre-service early childhood teachers' beliefs that influence their intention to use inquiry-based learning methods, *International Journal of Early Years Education.* https://doi.org/10.1080/09669760.2021.1890552

Education of Elementary Teachers on The Use of Adaptive Gamification Environment in Science Education

Alkinoos Ioannis Zourmpakis[1], Michail Kalogiannakis[2], Stamatis Papadakis[1]

[1] Department of Preschool Education, Faculty of Education, University of Crete, Greece

[2] Department of Special Education, University of Thessaly, Greece

This study investigates teachers' training regarding using an adaptive gamification application in teaching science education concepts in elementary education. Specifically, in this study, teachers are trained based on the Technological Pedagogical Science Knowledge (TPASK) model to utilize an adaptive gamification environment following a specific science education methodology. Our point of interest was to collect data on the teachers' perceptions regarding their readiness, self-efficacy, and motivation to integrate this digital application into the classroom. The data were collected through questionnaires before and after the teachers' training held remotely from December 2022 to early January 2023. They indicate an above-average perception regarding their TPASK dimensions, the increase of mainly the TPASK-specific dimension, and the enhancement of their understanding of the adaptive gamification environment and its use in teaching. Furthermore, teachers showed increased self-efficacy and motivation to use the application after the training, regardless of the training. The implication of this study indicates the need for similar training sessions regarding teachers' readiness to utilize adaptive learning digital applications and the acceptance and willingness to implement them in science education. They also provide valuable information for comparison during their use in the classroom.

Keywords: Adaptive gamification, teacher education, tpask framework.

Introduction

In recent years, technology has opened the path for a shift in education from static; one size fits models (one size fits all) to active, personalized and adaptive learning models (personal and adaptive learning) (O'Bannon & Thomas, 2015). Technology has the potential to provide a motivating learning experience that supports student-centred and adaptive learning models (Peng & Spector, 2019). For instance, adaptive gamification, i.e., adapting and adopting different playing

mechanisms and characteristic games based on the actions, preferences, and characteristics of each user/student, can accommodate the needs and interests of modern learners (Hassan et al., 2021). Especially in science education, the difficulty of comprehending and understanding physics contributes to negative emotions, experiences, and decreased motivation for learning (Brigido et al., 2013). However, though teachers recognize the value of technology, they often need help using it as a learning tool (UNESCO, 2017). Therefore, for teachers to effectively integrate ICT in education concerning the new pedagogical changes, they should come into contact and be trained with new models, which they should apply (Zourmpakis et al., 2022). For this purpose, an adaptive gamification application was created that fully integrates two learning strategies (problem-solving and inquiry-based learning) and multiple game elements (Zourmpakis et al., 2023).

Additionally, based on the Technological Pedagogical Science Knowledge (TPASK) model as a design framework (Jimoyiannis, 2010), a training program was developed for elementary in-service teachers to prepare them to teach scientific concepts with the use of an adaptive gamification environment in the actual classroom setting. Despite its apparent shortcomings, this study aims to provide data on the teachers' perceptions of readiness, self-efficacy, and willingness to integrate adapted gamification environments into science education lessons after a teacher training program. After conducting the training of the in-service teachers and analyzing the results of questionnaires, we will answer the following research question:

RQ: What are teachers' readiness, self-efficacy, and motivation to integrate adapted gamification environments into science education lessons after following a teacher training program?

Theoretical Framework

As new technologies emerge, such as gamification (Kalogiannakis et al., 2021), more and more researchers are investigating these tools to support scientific inquiry activities and address challenges related to inquiry learning. These challenges include insufficient student engagement, lack of necessary teacher guidance, and uncertainties in implementing inquiry activities in science education (Hugerat & Kortam, 2014; Tsai, 2018). Despite a decade of using gamification, i.e. employing game mechanics, aesthetics, and thinking related to games to engage and motivate students in order to facilitate learning and solve problems (Kapp, 2012), it has become clear that a 'one size fits all' approach is insufficient (Jagušt et al., 2018). Failure to tailor individual game elements and the instructional approach to the unique needs of each student, coupled with frequent presentations of similar and repetitive game elements, leads to increased dropout rates in the long term (Hassan et al., 2021). Users

often exhibit different playing styles and preferences in completing tasks and goals, typically motivated by specific game elements and mechanisms (Böckle et al., 2014). It is crucial to consider these factors during the design phase of a gamified application (Botte et al., 2020) and its implementation in the classroom (Zourmpakis et al., 2022).

The concept of adaptive gamification represents an innovative method of learning in which game elements are adapted based on the user's actions, preferences, and characteristics (Monterrat et al., 2017; Kalogiannakis et al., 2021). Adaptive gamification has been demonstrated to have varying effects on motivation, with positive outcomes such as increased engagement and retention (Hassan et al., 2021; Mora et al., 2018). However, this may only sometimes be the case, with some cases suggesting that both adaptive and non-adaptive elements can lead to similar experiences for students (Hallifax et al., 2019). Nevertheless, conflicting results in studies have raised concerns about user modelling and selecting relevant game elements (Hallifax et al., 2019). Although the field is still in its early stages of development (Zourmpakis et al., 2023), the shift in education from static and one-size-fits-all approaches to dynamic, personalized and adaptive learning methods (Courcier, 2012) has led to significant changes in the role of teachers. In the current landscape, teachers are tasked with providing strategic support to students, helping them to become self-directed and self-regulated learners capable of shaping their learning approaches (UNESCO, 2008).

Designing adaptive gamified environments is challenging (Seaborn & Fels, 2015), as game systems with an ideal range of game elements have yet to be found. Additionally, existing adaptive gamification approaches are primarily created generically (Monterrat et al., 2017). Nevertheless, the proper design of an adaptive gamified environment and an appropriate pedagogical approach are vital to increasing learner effectiveness and engagement in the long term (Böckle et al., 2017). The structured framework was based on two key elements. Firstly, the player model uses the HEXAD model (Tondello et al., 2019) to classify students' preferences into six primary categories influencing game elements and mechanics. The second element focuses on learning strategies, which are crucial in shaping the goals, objectives, game style and stages of the learning process. To optimize adaptation and familiarize learners, the framework highlights two learning strategies that share common aspects.

Furthermore, despite acknowledging the significance of technology, teachers frequently require assistance in utilizing it as a practical learning tool (UNESCO, 2017). Educators must understand the link between the integration of pedagogy and the use of digital technologies, especially in tools designed to support the entire learning process to foster the development of students as self-directed learners. Successful integration of Information and Communication Technology (ICT) requires teachers to be familiar with and trained in contemporary

models, such as the Technological Pedagogical Content Knowledge (TPACK) framework introduced by Mishra and Koehler (2006). In 2010, the concept of TPASK emerged, based on an integrated framework that draws on the theoretical principles of the TPACK model, authentic learning and expertise in teaching physical education (PE) (Jimogiannis, 2010). TPASK serves as a framework related to TPACK to equip teachers with the knowledge and skills necessary to teach science by integrating information and communication technology (ICT). This approach clarifies the components and establishes links between science education (content), pedagogy, and technology within a meaningful and realistic framework. By addressing the needs of teachers regarding the teaching of science education and the integration of ICT in the classroom, TPASK provides a comprehensive and practical basis for teacher preparation.

Methods & Context

The main interest of this study was to educate in-service teachers to improve their TPASK relationship, self-efficacy, and motivation to enhance their proficiency, ability, and willingness to apply their TPASK knowledge and skills regarding a digital application they have not used before. Table 1 provides some background information on the 7 participants that took part in it.

Table 1. Participants

Identification	Gender identification	Age group	Educational degree	Teaching experience	Self-assessment of ICT knowledge level
F1	Female	31-40	Masters	6-10	Good
M1	Male	22-30	Bachelor	0-5	Good
F2	Female	22-30	Masters	0-5	Good
F3	Female	51-60	Bachelor	21+	Fair
F4	Female	22-30	Masters	0-5	Fair
F5	Female	41-50	Bachelor	16-20	Good

This study was conducted in Heraklion, Crete, Greece, from December 2022 to early January 2023. National and international guidelines for research ethics were followed (Petousi & Sifaki, 2020). The data were collected through the use of questionnaires that were digitally distributed. The educational training was carried out remotely. Initially, we used a multiple-case experimental design with a convenience sample. However, as we examined each case study individually, we also incorporated a quantitative research approach, focusing on pre-and post-test measures despite the constraints of a limited sample size.

As Yazan (2015) emphasized, quantitative methods can contribute to multi-case

research studies.

Moreover, we followed a robust ethical protocol through the following: (a) external researchers checked questionnaires, (b) the questionnaires were pilot used with other teachers to check the questionnaire's clarity and inclusiveness, (c) inverted questions, (d) all data could be accessed and checked by external researchers to ensure the validity of this study (e) data were analyzed through the use of SPSS. This method was used for the following reasons: (1) it could give a clear and unbiased picture of the state of the in-service teachers before they utilize the digital application in the classroom and after (2). Additionally, many teachers were reluctant to participate in research requiring long hours of participation and thorough examination (3) The data could be cross-referenced with the TPACK Observation Checklist, which will be used during the classroom setting, and (4) interviews. However, this data cannot be generalized due to the small sample.

Findings

Results on TPASK

First, we will present each case participant individually. This paper will focus on the TPASK dimensions, where the participants reported an increase after training. Based on the participants' scores, none scored lower on their TPASK questionnaire after the training. Thus, an increase in the knowledge domain would be indicated if they noted an improvement in half of the questions. Subsequently, we will present the outcomes derived from a statistical analysis.

Regarding the F1 participant, she reported an average increase in the TK, PK, SK, TSK and TPASK knowledge. More specifically, they have acknowledged to be more proficient in the use of gamification applications (TK2) and adaptive gamification applications (TK4), utilizing a variety of learning strategies (PK1), various evaluation methods and techniques (PK3) and being able to handle adequately their classroom(PK2). Additionally, their scientific knowledge increased by following the evolution of science, particularly as it relates to the water cycle (CK2) and using the latest reference sources to increase their knowledge (CK3). Moreover, their TSK and TPASK knowledge improved as they were more qualified in using the technology to help understand the concepts of the water cycle (TCK1) or simulations associated with these concepts (TCK2), choosing learning strategies and technology that match the water cycle material to be delivered in class (TPASK2), integrating content knowledge, pedagogical knowledge, and technology knowledge to carry out meaningful learning (TPASK3) and enabling students to create deep and meaningful knowledge about water cycle material using adaptive gamification applications (TPASK5).

Regarding the M1 participant, he noted a general enhancement solely in the

enhancement of his TK. Specifically, he became more adept at following the latest technological developments proficiency (TK1) and utilizing adaptive gamification applications (TK4). Although there was no development in the majority of M1 participants' knowledge, it should be noted that the average score in all knowledge parts was higher than the average in pre-tests. Though his knowledge in the TPASK domains did not improve generally after the training, it stayed satisfactory.

The F2 participant was able to develop her knowledge of TK, TPK, PSK, TSK, and TPASK. Admittedly, she became more capable in the use of gamification applications (TK2), simulation applications (TK3), and adaptive gamification applications (TK4). Also, she has identified as being more skilled in implementing gamification applications to communicate with students (TPK3) and gamification platforms to do tests (TPK4). In addition, her pedagogical science knowledge and technological science knowledge were enhanced as they became more competent in assisting students to understand concepts (PSK1) further, actively solving real-world problems (PSK3), changing their thinking process to master a complex topic related to water cycle concepts without the use of technology (PSK4), develop student activities and projects that involve the use of gamification (TSK3) and employ adaptive gamification applications related to water cycle concepts (TSK4). Moreover, she acknowledged that she became pretty more skilled in integrating scientific, pedagogical and technological knowledge (TPASK1), learning strategies and technologies appropriate to the water cycle material to be presented in the classroom (TPASK2), content knowledge, pedagogical knowledge and technological knowledge to deliver meaningful learning (TPASK3), different learning strategies and different technologies to the application of science learning (TPASK4) and facilitating students to create deep and meaningful knowledge about the material of the water cycle using adaptive gamification applications (TPASK5).

Regarding the F3 participant, she noticed an improvement in her TSK and TPASK knowledge. Specifically, she became more proficient in employing simulations (TSK2) and adaptive gamification applications related to water cycle concepts (TSK4). Her TPASK developed significantly as she improved her ability to blend scientific, pedagogical and technological knowledge (TPASK1), to combine content knowledge, pedagogical expertise and technological expertise to deliver learning effectively (TPASK3), to use different learning strategies and technologies for science teaching (TPASK4) and to guide students to build a deep understanding of water cycle material through adaptive gamification applications (TPASK5).

Regarding the F4 participant, she reported higher growth in their SK, TPK and TSK knowledge. She believed that she could comprehend the concepts related to the water cycle, such as melting, freezing, boiling and evaporation

(SK1), and keep up to date with scientific advances, particularly about the water cycle (SK2). Furthermore, she indicated that she could make use of computer applications in every lesson plan (TPK1), select technology that fits her pedagogical approach and learning strategy in her lessons (TPK2) and employ gamification applications to engage with students (TPK3). Her technological science knowledge displayed improvement as she noted having the capability to design student activities and projects incorporating gamification (TSK3) and employing adaptive gamification applications linked to concepts within the water cycle (TSK4).

In the last participant, F5, a noticeable positive enhancement in TK, PSK, TSK and TPASK was observed. She noted that she gained increased competency in using simulation (TK3) and adaptive gamification environments (TK4). Additionally, her PSK was significantly boosted as she reported having the ability to assist students in their grasp of water cycle concepts through a variety of strategies (PSK1), actively engage with real-world problems related to water cycle concepts (PSK3), guide the transformation of students' thought processes to master challenging aspects related to water cycle concepts (PSK4) and encourage meaningful student reflection on water cycle concepts without relying on technology (PSK5). Moreover, she was able to harness technology to support understanding of water cycle concepts (TSK1), apply simulations related to water cycle principles (TSK2), create student activities and projects that integrate gamification (TSK3), and implement adaptive gamification applications linked to water cycle concepts (TSK4). Furthermore, She recognized a significant improvement in her ability to integrate scientific, pedagogical, and technological knowledge (TPASK1), employ appropriate learning strategies and technologies for presenting water cycle material in the classroom (TPASK2), combine content knowledge pedagogical expertise, use different learning strategies and technologies in science education applications (TPASK4), and guide students in constructing a profound understanding of water cycle material through the use of adaptive gamification applications (TPASK5).

Based on our statistical analysis (Table 2), the in-service teachers rated most of their TPASK dimensions higher than neutral before the educational intervention, with the exceptions being slightly below. After we conducted the educational process, the teachers improved most of the TPASK dimensions. However, a statistically significant difference was mainly observed in the TPASK dimension. Specifically, the teachers appeared to benefit in all aspects of the TPASK but showed significant improvement in those associated with using adaptive gamification. As such, their post questionnaires regarding their readiness based on their TPASK showed significantly higher results regarding the application use.

Results on Self-Efficacy, Perceived Usefulness, Perceived Ease of Use, Attitudes Towards Computer Use and Behavioural Intentions

Similar to the presentation of the TPASK results, the results for self-efficacy (SE), perceived usefulness (PU), perceived ease of use (PEU), attitudes towards computer use (ATCU) and behavioural intentions (BI) are presented first on an individual case participant basis. In contrast to the TPASK results, some teachers showed a decrease in their post-test scores. More specifically, this decrease was in the perceived usefulness questions. However, this decrease will be discussed outside of this paper. On the other hand, all other dimensions either increased or remained the same. Therefore, an increase in the knowledge domain would be suggested if participants observed an improvement in at least half of the questions. We will then also present the results based on statistical analysis.

Regarding the F1, she displayed an increase solely in her self-efficacy. More specifically, her confidence was raised to teach her students using adaptive gamification applications (SE2), is more able to overcome all challenges she faces to achieve her teaching goals (SE6), take full advantage of technological innovations (SE7) and engage even the most reluctant and challenging students in her classroom activities using adaptive gamification applications (SE8). Additionally, she was more able to organize and complete her work even when faced with unexpected or challenging tasks, deal effectively with her students' problem behaviours and gain the trust and respect of all her colleagues.

The M1 participant reported higher scores in his self-efficacy and perceived ease of use. Notably, he had higher confidence levels in teaching his students using adaptive gamification applications (SE2) and training her colleagues on using adaptive gamification apps in education (SE3). What is more, he believed that he was more adept at using various technological tools in his classroom without anyone showing him (SE5), fully utilizing technological innovations in his teaching (SE7), engaging even the most reluctant and challenging students in his classroom activities using adaptive gamification applications (SE8), and organizing and completing their work even when faced with unexpected or challenging tasks (SE9). Also, it was easier for him to do tasks he wanted to do using adaptive gamification apps (PEU1), and his interaction with the gamification application was clear and understandable (PEU3).

The F2 participants' self-efficacy, perceived ease of use, attitude toward computer use, and behavioural intention were enhanced. Specifically, she feels more comfortable implementing adaptive gamification in her classroom independently (SE1) and feels confident teaching her students using gamification apps (SE2). She is also confident that she can effectively train her colleagues to use these educational applications (SE3). Using technology tools in her classroom is a seamless task (SE4), and she can easily navigate them independently without guidance (SE5). In addition, she can use technological innovation in her teaching

(SE5), successfully engaging even the most reluctant students through adaptive gamification applications (SE8), and the ability to organize and complete work efficiently, even when faced with unforeseen or complex tasks (SE9).

Moreover, she indicated that she can effortlessly perform tasks by employing adaptive gamification apps (PEU1); utilizing adaptive gamification applications is a straightforward (PEU2) and enjoyable experience for her (PEU3), making the learning process more interesting (ATCU1) and fun (ATCU2). Furthermore, she expressed a more substantial commitment to integrating adaptive gamification applications for teaching and learning in science education (BI1), with an increased determination to use them whenever possible (BI2). In addition, she is inclined to increase the frequency of using adaptive gamification applications in the future (BI3). She remains steadfast in her commitment to their regular integration into her teaching methods (BI4).

Regarding the F3 participant, her notable change was primarily evident in the increased perception of ease of use and a more positive attitude toward computer use. To be more precise, she conveyed that using adaptive gamification applications makes it easy for her to complete tasks (PEU1). In addition, she found the use of adaptive gamification applications easy (PEU2) and fun (PEU3) and enhanced the learning process by adding interest (ATCU1) and fun (ATCU2). It also shows a positive tendency to use adaptive gamification applications (ATCU3).

Participant F4 demonstrated improved self-efficacy, perceived usefulness, ease of use, and attitude toward computer use. As she reported, her self-efficacy has notably risen, as indicated by her comfort in using adaptive gamification independently in the classroom (SE1), confidence in teaching with adaptive gamification apps (SE2), and assurance in training colleagues on their use in education (SE3). Also, she demonstrated increased proficiency in operating technology tools (SE4), even independently (SE5) and overcoming challenges to achieve teaching goals (SE6). She exhibited a notable enhancement in her ability to make full use of technological innovations in teaching (SE8), organize and complete work even when faced with unexpected or demanding tasks (SE9), her effectiveness in dealing with students' challenging behaviour (SE10), and her success in gaining the trust and respect of all her colleagues (SE11). Her perception of tasks involving these applications as easy to execute also experienced a notable increase (PEU1, PEU2), and her interaction with computers became more transparent and more understandable (PEU3). In terms of attitude toward computer use, she significantly increased her appreciation, recognizing that using adaptive gamification applications enhances the learning process, making it more exciting and enjoyable (ATCU1, ATCU2). She also expressed a heightened positive liking for utilizing adaptive gamification applications (ATCU3).

As for Participant F5, she exhibited a modest increase in her self-efficacy. Notably, she expressed a heightened comfort in using adaptive gamification independently in her classroom (SE1). Also, her confidence in teaching students with adaptive gamification apps (SE2) and training colleagues on their use in education (SE3) has significantly grown. Furthermore, she indicated an increased ability to take full advantage of technological innovations in her teaching (SE7), and she is now better able to engage even the most reluctant and challenging students through adaptive gamification applications (SE8).

Regarding the teachers' self-efficacy, they displayed a statistically significant difference in almost half the questions of this dimension, especially concerning the use of adaptive gamification applications in the classroom. Additionally, although there was no statistically significant distinction, teachers displayed a high average before and after the training on their intention to use this application in their teachings. Moreover, the teachers appeared motivated to apply what they had learned in their teachings but needed more motivation to share this with their colleagues. Given the page limitation, Table 2 consists only of the questions that showed significant differences in how they were coded.

Table 2. Questions with a significant difference

Dimension	Question	Mean difference	Std Deviation	Significant difference (sd<,05)
Technological Knowledge (TK)	TK4	-,571	,535	,030
Pedagogical Knowledge (CK)	CK2	-1,000	,577	,004
Technological Content Knowledge (TCK)	TCK4	-,857	1,069	,078
TPASK	TPASK1	-,714	,756	,047
TPASK	TPASK2	-,429	,535	,078
Self-efficacy	SE2	-,857	,900	,045
Self-efficacy	SE3	-1,143	,900	,015
Self-efficacy	SE7	-1,143	,378	,000
Self-efficacy	SE8	-1,000	,577	,004
Self-efficacy	SE9	-1,000	1,000	,038
Perceived Ease of Use	PEU3	-,571	,535	,030
Attitude Toward Computer Use	ATCU2	-,714	,286	,047

Discussion

Following the implementation of the training programme, significant improvements were observed in specific dimensions of technological, pedagogical and scientific knowledge (TPASK). As shown in each case and the

statistical analysis, preservice teachers showed some improvement in TK, CK and TSK, primarily related to using adaptive gamification and science content. This aligns with prior research findings indicating that preservice teachers typically possess more advanced Content Knowledge (CK) and Pedagogical Knowledge (PK) when compared to other knowledge domains (Doering et al., 2014; Koh & Chai, 2014). However, preservice teachers showed a significant improvement in their TPASK knowledge, consistent with the findings of other studies (Jimogiannis, 2010; Koh & Divaharan, 2011). In addition, it was shown that most of the improvement in TPASK knowledge occurred among women. As Huang and Fraser (2008) noted, female science teachers are more interested in their profession than their male counterparts. This may explain the greater attention and confidence in developing their TPACK perceived by the female preservice teachers in this study. A statistically significant difference was observed in the aspect related to the use of adaptive gamification.

Regarding teacher self-efficacy, almost all teachers expressed an increase in self-efficacy. Statistical analysis supported this, particularly by incorporating adaptive gamification applications into the educational environment. The study results suggest that the intervention is promising for increasing preservice teachers' self-efficacy, consistent with previous research findings (Cetin-Dindar et al., 2018; Zimmerman et al., 2021). Previous studies have found a direct correlation between teachers' self-efficacy and their technological pedagogical content knowledge (TPACK) (Wang & Zhao, 2021; Yildiz Durak et al., 2023). Educators with higher levels of self-efficacy are more likely to experience positive emotions when integrating information and communication technologies (ICT) in the classroom (Yildiz Durak et al., 2022). Therefore, it is crucial to support their professional development through pedagogical training programmes that enable them to incorporate digital technologies in relevant contexts, develop appropriate lesson plans, and implement real classroom scenarios to enhance student learning in specific subject areas (Roussinos & Jimoyiannis, 2019).

Conclusions

As highlighted, although showing above-average readiness based on the TPASK dimensions, self-efficacy and motivation to use new digital applications, in-service teachers can benefit from training programmes with the TPASK approach, especially in using and integrating these innovative digital applications in their teaching. These findings are significant because they showcase the acceptance and willingness of teachers to integrate digital applications that encompass active personalized and adaptive learning models in teaching science education and the need to develop educational training programs with the TPASK approach. Nonetheless, the results of this study can not be generalized as the survey sample is tiny.

Teachers accept and support this transition as education moves towards more student-centred and personalized learning environments using digital technologies. They can benefit from similar training programmes, especially in using and integrating these innovative digital applications into their teaching. However, their willingness, motivation and perceived readiness do not guarantee their success in the following classroom implementation. Furthermore, the implementation in the classroom and the interviews that would take place afterwards will help to point out additional aspects and, through triangulation, validate our findings.

References

Brígido, M., Borrachero, AB, Bermejo, ML, & Mellado, V. (2013). Prospective primary teachers' self-efficacy and emotions in science teaching. *European journal of teacher education, 36* (2), 200-217. https://doi.org/10.1080/02619768.2012.686993

Cetin-Dindar, A., Boz, Y., Sonmez, D. Y., & Celep, N. D. (2018). Development of preservice chemistry teachers' technological pedagogical content knowledge. *Chemistry Education Research and Practice*, *19*(1), 167-183. https://doi.org/10.1039/C7RP00175D

Codish, D., & Ravid, G. (2014). Academic course gamification: The art of perceived playfulness. *Interdisciplinary Journal of E-Learning and Learning Objects*, *10*(1), 131–151.

Courcier, I. (2012). What is Personalized Learning in England? An Investigation of Teacher's Conceptions. In Personalisation of Education in Concepts (pp. 141–160). SensePublishers.

Hallifax, S., Serna, A., Marty, J. C., & Lavoué, É. (2019). Adaptive gamification in education: A literature review of current trends and developments. In *Transforming Learning with Meaningful Technologies: 14th European Conference on Technology Enhanced Learning, EC-TEL 2019, Delft, The Netherlands, September 16–19, 2019, Proceedings 14* (pp. 294-307). Springer International Publishing.

Hassan, M. A., Habiba, U., Majeed, F., & Shoaib, M. (2021). Adaptive gamification in e-learning based on students' learning styles. *Interactive Learning Environments*, *29*(4), 545-565. https://doi.org/10.1080/10494820.2019.1588745

Huang, S. Y. L., & Fraser, B. J. (2009). Science teachers' perceptions of the school environment: Gender differences. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *46*(4), 404–420. https://doi.org/10.1002/tea.20284

Hugerat, M., & Kortam, N. (2014). Improving higher order thinking skills among freshmen by teaching science through inquiry. *Eurasia Journal of Mathematics, Science and Technology Education*, *10*(5), 447-454. https://doi.org/10.12973/eurasia.2014.1107a

Jagušt, T., Botički, I., & So, H. J. (2018). Examining competitive, collaborative and adaptive gamification in young learners' math learning. *Computers & Education*, *125*, 444-457. https://doi.org/10.1016/j.compedu.2018.06.022

Jimoyiannis, A. (2010). Designing and implementing an integrated technological pedagogical science knowledge framework for science teachers professional development. *Computers & Education*, *55*(3), 1259-1269. https://doi.org/10.1016/j.compedu.2010.05.022

Kalogiannakis, M., Papadakis, S., & Zourmpakis, A.-I. (2021). Gamification in Science Education. A Systematic Review of the Literature. *Education Sciences*, *11*(1), 22. http://dx.doi.org/10.3390/educsci11010022

Kapp, K. M. (2012). *The gamification of learning and instruction: game-based methods and strategies for training and education*. John Wiley & Sons.

Koh, J. H., & Divaharan, H. (2011). Developing preservice teachers' technology integration expertise through the TPACK-developing instructional model. *Journal of Educational Computing Research*, *44*(1), 35-58. https://doi.org/10.2190/EC.44.1.c

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. Teachers College Record, 108(6), 1017-1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Monterrat, B., Lavoué, É., & George, S. (2017). Adaptation of gaming features for motivating learners. *Simulation & Gaming*, *48*(5), 625-656. https://doi.org/10.1177/1046878117712632

Mora, A., Tondello, G. F., Nacke, L. E., & Arnedo-Moreno, J. (2018). Effect of personalized gameful design on student engagement. In *2018 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1925-1933). IEEE. https://doi.org/10.1109/EDUCON.2018.8363471

O'Bannon, B. W., & Thomas, K. M. (2015). Mobile phones in the classroom: Preservice teachers answer the call. *Computers & Education*, *85*, 110–122. https://doi.org/10.1016/j.compedu.2015.02.010

Peng, H., Ma, S., & Spector, J. M. (2019). Personalized adaptive learning: an emerging pedagogical approach enabled by a smart learning environment. *Smart Learning Environments*, *6*(1), 1–14. https://doi.org/10.1186/s40561-019-0089-y

Petousi, V., & Sifaki, E. (2020). Contextualizing harm in the framework of research misconduct. Findings from discourse analysis of scientific publications, *International Journal of Sustainable Development, 23*(3/4), 149-174, https://dx.doi.org/10.1504/IJSD.2020.10037655

Roussinos, D., & Jimoyiannis, A. (2019). Examining primary education teachers' perceptions of TPACK and the related educational context factors. *Journal of Research on Technology in Education*, *51*(4), 377-397. https://doi.org/10.1080/15391523.2019.1666323

Tondello, G. F., Mora, A., Marczewski, A., & Nacke, L. E. (2019). Empirical validation of the gamification user types hexad scale in English and Spanish. *International Journal of Human-Computer Studies*, *127*, 95-111. https://doi.org/10.1016/j.

ijhcs.2018.10.002

Tsai, F. H. (2018). The development and evaluation of a computer-simulated science inquiry environment using gamified elements. *Journal of Educational Computing Research*, *56*(1), 3–22. https://doi.org/10.1177/0735633117705646

UNESCO. (2017). *UNESCO's In pursuit of smart learning environments for the 21st century.* https://unesdoc.unesco.org/ark:/48223/pf0000252335

UNESCO. (2008). *UNESCO's ICT Competency Standards for Teachers*. https://unesdoc.unesco.org/ark:/48223/pf0000156207.

Wang, Q., & Zhao, G. (2021). ICT self-efficacy mediates most effects of university ICT support on preservice teachers' TPACK: Evidence from three normal universities in China. *British Journal of Educational Technology*, *52*(6), 2319–2339. https://doi.org/10.1111/bjet.13141

Yildiz Durak, H., Atman Uslu, N., Canbazoğlu Bilici, S., & Güler, B. (2023). Examining the predictors of TPACK for integrated STEM: Science teaching self-efficacy, computational thinking, and design thinking. *Education and Information Technologies*, *28*(7), 7927-7954. https://doi.org/10.1007/s10639-022-11505-7

Zimmermann, F., Melle, I., & Huwer, J. (2021). Developing prospective chemistry teachers' TPACK–A comparison between students of two different universities and expertise levels regarding their TPACK self-efficacy, attitude, and lesson planning competence. *Journal of Chemical Education*, *98*(6), 1863-1874. https://doi.org/10.1021/acs.jchemed.0c01296

Zourmpakis, A. I., Papadakis, S., & Kalogiannakis, M. (2022). Education of preschool and elementary teachers on the use of adaptive gamification in science education. *International Journal of Technology Enhanced Learning, 14*(1), 1-16. https://doi.org/10.1504/IJTEL.2022.120556

Zourmpakis, A. I., Kalogiannakis, M., & Papadakis, S. (2023). A Review of the Literature for Designing and Developing a Framework for Adaptive Gamification in Physics Education. in Taşar, M. F.& Heron, P. R. L. (Eds.), *The International Handbook of Physics Education Research: Teaching Physics* (pp. 5–1–5–26) AIP Publishing. https://doi.org/10.1063/9780735425712_005.

Development and Implementation of A Robotic Activity Through Bbc Micro:Bit for Teaching Electricity to 5th Graders

Effransia Tzagkaraki[1], Michail Kalogiannakis[2] and Stamatios Papadakis[1]

[1] Department of Preschool Education, University of Crete, Greece

[2] Department of Special Education, University of Thessaly, Greece

This research examines whether teaching using the proposed application of educational robotics improves students' school performance to a greater extent than traditional teaching in electricity for 5th-grade students. It is also examined whether their attitude towards Physics is affected after the implementation of the activity. A Teaching Learning Sequence (TLS) was designed and implemented to achieve this aim. A quasi-experimental design with a pre-test and post-test was applied for 140 participants. Two questionnaires were used as research tools. As it turned out, the experimental group performed better overall. Children from both groups were not significantly affected in their view of the importance of Physics. However, they showed an increased desire to engage in it.

Keywords: Robotics, Micro:bit, electricity.

Introduction

Science education refers to teaching and learning concepts, theories, and the scientific method that will help to understand the world and the universe beyond it. It is therefore expected that by utilizing conceptual knowledge and skills, students will be able to solve problems in their environment (National Research Council, 2007). In a traditional approach to teaching Physics, teachers, mainly through lectures, present the learning content, which students then apply to experimental designs and problems. However, in such an approach, knowledge appears fragmented, and this gap is filled by the holistic approach advocated by STEM education (Chaldi & Mantzanidou, 2021; Foti, 2021). Combining knowledge and skills from various fields makes students more likely to respond successfully to various situations (Chiu, Hwang & Tu, 2022).

Several researchers have reported that engaging with educational robotics can become the means to design and implement a variety of activities and effectively contribute to learning and skill enhancement (Kalogiannakis & Papadakis, 2022;

Tzagaraki, Papadakis, & Kalogiannakis, 2022; Tzimopoulos, Provelengios, & Iosifidou, 2021). Primarily for physics, through experience, abstract concepts of physics and other fields can be understood while simultaneously promoting cooperation and communication skills. In addition, robotic activities can effectively engage children with Inquiry-Based Learning, where the scientific way of working is followed (Kyriazopoulos, Koutromanos, Voudouri, & Galani, 2021).

Theoretical Framework

The Use Of BBC Micro:Bit

In such a context, physical and visual programming through physical devices (such as the BBC Micro:bit) includes various projects' design and development processes. The use of physical devices can be a tool used originally and creatively in different educational situations and in various educational applications where students can develop their ideas and initiatives interactively for collaborative and individual projects (Vordou & Romero, 2021) and a Teaching Learning Sequence an appropriate method (Méheut & Psillos, 2004).

The BBC Micro:bit is an effective educational tool for promoting programming and fostering interest in various STEM subjects. The user can program this microprocessor to perform some essential functions, and by combining some of its features, he can develop innovative digital solutions (Kalogiannakis, Tzagkaraki & Papadakis, 2021).

The teacher can take different approaches to support students' learning through the Micro:bit. Clear goals, group, pair, or even individual work, provision of appropriate feedback along the way, and differentiated activities adapted to the needs and capabilities of children of this age contribute in this direction (Sentance, Waite, MacLeod & Yeomans, 2017).

Attitudes and Motivation' S Role

In various social-psychological models, there is a convergence in the view that the most direct and significant predictor of a person's behaviour is his intention to perform it (Sheeran, 2002). Understanding the attitudes towards science in education is essential, as it can affect how they function in the learning environment and their participation in related activities. Also, a relationship exists between students' attitudes and academic achievements, indirectly through motivation. Given that motivation is a critical factor related to the propensity to participate in science learning activities, lack of motivation has often been linked to low performance and limited prospects for employment in related science fields. To be able to examine the process of conceptual change for the applications of physical science (in the sense of changing children's perceptions and opinions, through physical programming in our case), the emotions they

experience when facing its challenges, the degree of engagement in achieving learning goals, it is necessary to study children's motivation as well as how we assess this motivation (You, Kim, Black & Min, 2018).

The TLS we developed is a short-term program to engage primary school students with physical and visual programming applications. In designing the learning activities, activities structured for whole-class participation and group activities based on smaller groups were used. The experimental group implemented these dynamic activities, focusing on the children.

Method

Aim – Research Questions

The purpose of the study is twofold: It examines whether teaching using the proposed application of educational robotics improves, to a greater extent, students' school performance compared to traditional teaching in electricity for 5th-grade students. It is also examined whether their attitude towards Physics is affected after the implementation of the activity.

The research questions of the study based on the purpose are:

RQ1: Did the children's performance improve in understanding concepts and applications of electricity?

RQ2: Is there an effect on the attitude of the children of the experimental group regarding the course of Physics after engaging with the topic of electricity through TLS?

Research Strategy – Sample – Research Tools

A quasi-experimental design was applied. The research was implemented in three phases from February to April 2023. Two questionnaires were used as research tools. One questionnaire concerns the investigation of their attitudes towards Physics, and the other concerns evaluating the children's performance.

The research tool for attitudes is a questionnaire adapted to the specific research needs (https://rb.gy/uqauw5) with six categories; each includes five questions except the first, which includes three. The categories are the following: Perceptions of the teacher, the importance of Physics to society, Intrinsic motivation, Career motivation, Self-efficacy, and Grade motivation. In April 2022, a pilot questionnaire was applied to 30 students, resulting in changes to its final form. Care has been taken to ensure the questions are worded so all children can respond. All required permits secured. A five-activity paper-based assessment was developed to assess students' knowledge of electrical concepts. Assignments given to students are available at the link: https://shorturl.at/filnR. We have considered the official syllabus of the specific country for the assessment

work. Each activity was evaluated with 6 points, and the total evaluation gave a maximum of 30 points. We also capture data from programming the board during the experimental process via the BBC Micro:bit platform.

In the first phase, students were given a questionnaire of 28 attitude detection questions and then the other with a five-activity printed questionnaire (Pre-test). In the second phase, the teaching intervention of 10 teaching hours was carried out with the implementation of the planned activities. After completing the experimental design, the Post-test was conducted to evaluate the effect of the experimental teaching, according to the research purpose, using the same questionnaires.

Firstly, students explored how Microsoft's MakeCode program works as it uses colour-coded blocks familiar to anyone who has used Scratch before and gives access to all the board's features. The simulation screen, command palettes, and code generation areas were used. The sequential and control structure was used from the basic programming structures. In the second part, students in groups of four (allocation was done via ClassDojo) implemented a scenario in real-time. The scenario concerned constructing a movement mechanism where a motor was used. Before the post-test, each group presented their project to the others. The Ethics and Research Ethics Committee of the University of Crete

reviewed and approved the research protocol.

The study sample was 140 students from four different Greek public schools. In each school, there were two different departments. One of the two was chosen by lottery as the experimental group, so the two groups (control and experimental) were equivalent. All students were born in 2012 and had a basic knowledge of visual programming commands. The majority had used Scratch for more than five teaching hours, as informed by the IT teacher. Informed consent was secured (Petousi & Sifaki, 2020).

Results

The following sections present the main results based on the student's answers and their statistical processing. We analyze each of the six categories of attitudes and motivations by focusing on individual questions (e.g. Q4, Q12 or Q26) or all questions related to a category (e.g. career motivation) through tables of descriptive statistics to facilitate understanding.

Investigating Attitudes

In investigating attitudes, six categories that are related to attitudes are examined (Figure 1) (Hillman, Zeeman, Tilburg, et al., 2016; Salta & Koulougliotis, 2015; Tai, Ryoo, Skeeles-Worley, et al., 2022). To check the reliability of the data, the Cronbach α index was calculated with a value of 0.879.

The experimental group in 5 of the six categories examined increased its agreement percentage on the questionnaire's relevant statements. A decrease in the percentage of agreement was noted for the category concerning intrinsic motivation (Figure 2).

Figure 1. Considered categories for the attitudes.

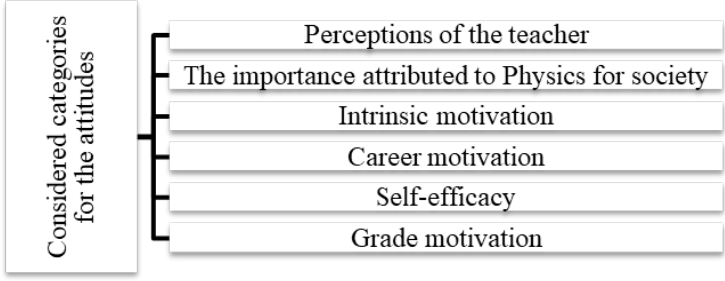
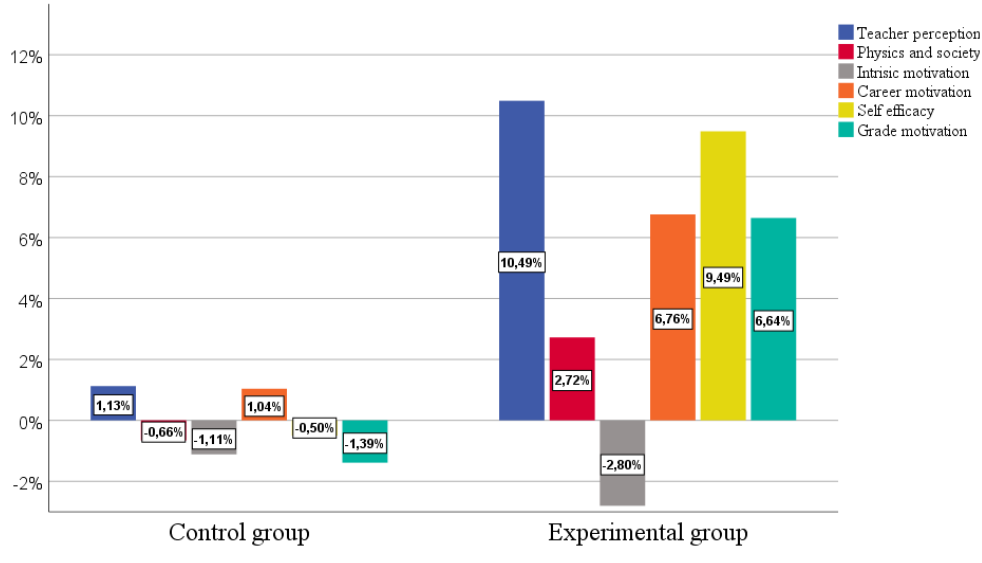


Figure 2. Post-test differences per category and group.

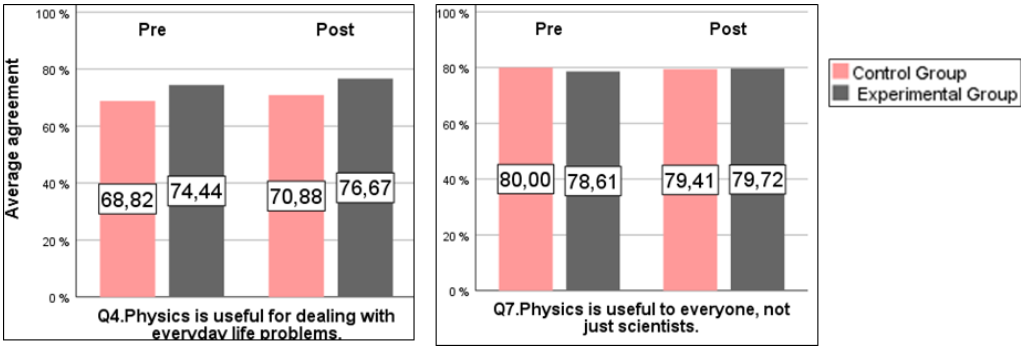


To the question of whether there are differences in the perceptions of the teacher, The students tend to agree that the way the teacher presents the material and his/her intention to offer individual help affects their interest in the lesson. Indicate the percentage of agreement of the experimental group on the statement: The physics teacher interestingly presents the material. 89% of the experimental group agree, and 75% of the control group at post-test.

For the experimental group, the attitude towards the contribution of physics to society increased by about 3%. Children state that physics contributes to solving

everyday problems and is beneficial not only for scientific purposes (Figure 3). In the second question, no significant differences are observed, but both groups' high percentages of agreement are remarkable.

Figure 3a, b. The average agreement of the two groups on the importance attributed to Physics for society on two indicative statements.



The experimental group noted the lowest percentage of agreement (-2.80%) for the intrinsic motivation category. They find physics class easy (Figure 4) but seem thoughtful about how what is being taught relates to their lives (Figure 5).

Figure 4a, b. Q9. Physics is easy for me.

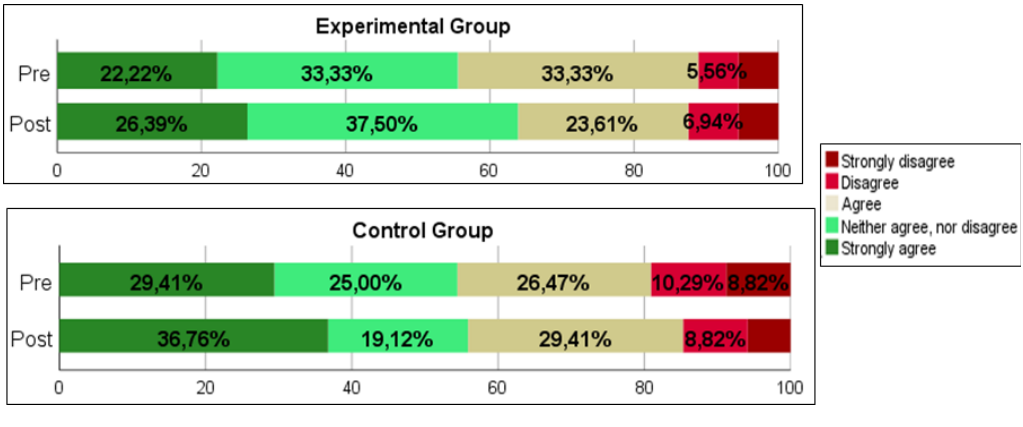
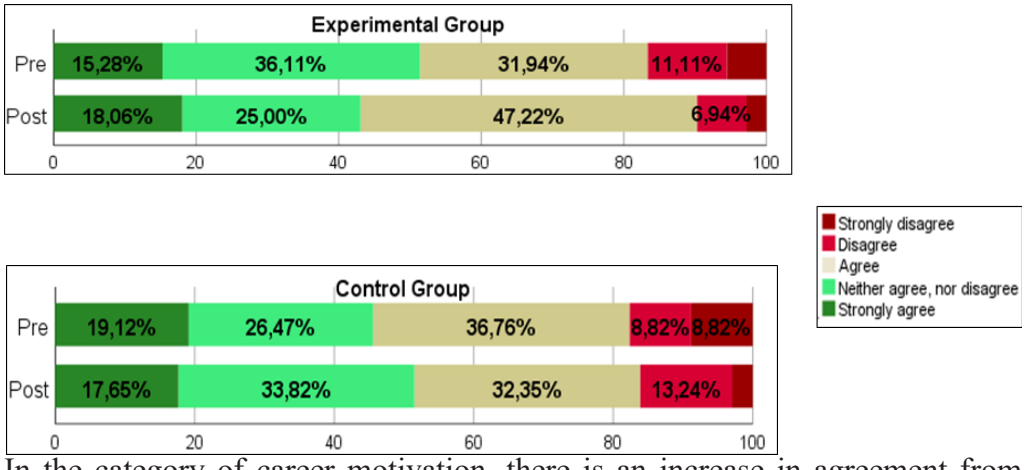


Figure 5a, b. Q12. The Physics I learned is related to my life.



In the category of career motivation, there is an increase in agreement from the experimental group of 7% in the post-test. Regarding career motivation, we observed an increase in the agreement rate of students in the experimental group on whether they would like a job in the future related to physics and whether it would benefit their career. In the question about their future profession and whether they would like it to be related to physics, it is observed that the experimental group increased their agreement by about 15%. In comparison, the control group decreased by 19%. Table 1 shows the descriptive statistics for this category.

Table 1. Descriptive statistics for Career motivation category.

Career Motivation Mean		Control Group		Experimental Group	
		Std. Dev.	Mean	Std. Dev.	
Q14. I want a job related to Physics because I am interested.	Pre	3.16	1.253	2.35	1.212
	Post	2.71	1.185	2.75	1.172
Q15. I will use physics problem-solving skills in my work.	Pre	3.13	1.183	3.22	1.010
	Post	3.06	1.220	3.04	1.156
Q16. Learning Physics will give me a professional edge.	Pre	3.35	.989	3.32	1.173
	Post	3.44	1.164	3.19	1.030
Q17. Understanding Physics will benefit my career.	Pre	3.63	1.145	3.43	1.136
	Post	3.56	.998	3.61	.943
Q18. Learning Physics will help me get a good job.	Pre	3.24	1.186	3.13	1.034
	Post	3.40	1.223	3.04	1.080

In the category of self-efficacy, the mean is 3.46 for the control group and 3.57 for the experimental group at the post-test. The difference in the means between the groups is insignificant; however, the percentages of agreement increase for

the experimental group and decrease in most statements for the control group. More specifically, understanding was improved for the experimental group. In Q20, the control group showed the most significant increase (5.90%) in their engagement time in understanding the Physics concepts being taught. The experimental group's perception that they would respond well to the physics assignments was enhanced by 2.50%, too. (Table 2).

Table 2. Descriptive statistics for the Self-Efficacy category.

Self-Efficacy Mean		Control Group		Experimental Group	
		Std. Dev.	Mean	Std. Dev.	
Q19. I understand what we are talking about in Physics.	Pre	3.94	.944	3.69	1.182
	Post	3.66	1.167	3.90	1.023
Q20. I spend much time trying to understand physics.	Pre	2.75	1.460	2.67	1.424
	Post	3.04	1.332	2.76	1.477
Q21. I believe I can get an "excellent" in Physics.	Pre	3.81	1.284	3.78	1.366
	Post	3.51	1.252	3.85	1.109
Q22. I will do well in the Physics assignments.	Pre	3.72	1.256	3.65	1.103
	Post	3.57	1.250	3.78	1.103
Q23. I will do well in the Physics exams.	Pre	3.63	1.078	3.40	1.252
	Post	3.53	1.190	3.54	1.150

Figure 6a, b. The average agreement of the two groups on the importance attributed to Grade motivation on two indicative questions.



Grade motivation increased for the experimental group by approximately 7% concerning the control group. Good grades in Physics are essential for both groups. We can see Q24 and Q26 combined. For Q24, 82.00% of the control group and 86.40% of the experimental group in the pre-test consider it essential to do well in Physics. The experimental group in the post-control group maintained this percentage, while a decrease of about 6% was observed in the

control group. For Q26 at post-test, both groups increased their percentage of agreement (Figure 6).

Students School Performance

The non-parametric Wilcoxon test (Table 3) was chosen to identify whether there is a statistically significant difference between the scores. As can be seen, most of the participants from both groups had a higher score in the post-test than in the pre-test.

Table 3. Wilcoxon Signed Ranks Test.

		N	Mean Rank	Sum of Ranks
Total post – Total pre	Negative Ranks	2[a]	8,50	17,00
	Positive Ranks	137[b]	70,90	9713,00
	Ties	1[c]		
	Total	140		

a. Total post < Total pre

b. Total post > Total pre

c. Total post = Total pre

A statistically significant difference between the two methods is identified through a Dependent Samples t-test. The proposed method is effective as the children's performance improved in the post-test. Therefore, the hypothesis of no differentiation between the two methods is rejected.

Table 4a, b. Dependent Samples t-test.

Paired Samples Statistics				
	Mean	N	Std. Dev.	Std. Error Mean
Total pre	56,4524	140	16,63647	1,40604
Total post	85,6429	140	12,46512	1,05349

				95% Confidence Interval of the Difference				
	Mean	Std. Dev.	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Total pre – Total post	-29,190	15,928	1,346	-31,852	-26,529	-21,684	139	,000

Examining the students' performance, the performance improvement rates were 27.04% for the control group and 30.28% for the experimental group. In the individual questions, no significant differences were observed. Although the experimental group generally performed slightly better. The table below (Table 5) shows each group's improvement at the post-test for each of the five tasks of

the tool. All tasks were equivalent, and the maximum score was 6 points.

Table 5. Improvement per task.

Task	Control group	Experimental group	Difference between groups
1	35%	56%	21%
2	31%	53%	22%
3	170%	157%	-13%
4	40%	36%	-4%
5	20%	16%	-4%

Discussion -Conclusion

This paper explores learning electricity concepts through physical block-based programming and analyses children's experiences as reflected in their questionnaire responses. In terms of attitudes, six categories were examined. The teacher and how the material is presented influence the group's interest. This is also linked to the conclusion of various research where engaging with physical programming and the Micro:bit stimulates children's interest as it is a more fun way of teaching (Przybylla & Romeike, 2018), and it helps them experience learning as an invention rather than imitation or just implementation of algorithms (Min & Kim, 2020).

The teacher and how the material is presented influence the group's interest. In particular, the agreement rate for the experimental group regarding the manner of presentation increased by 10% (from 77% to 87% in the post-test). This is also linked to the conclusion of various research where engaging with physical programming and the Micro:bit stimulates children's interest as it is a more fun way of teaching (Przybylla & Romeike, 2018) and helps them experience learning as an invention instead of imitation or simply implementing algorithms (Min & Kim, 2020).

Children in the experimental group have slightly increased their career-related motivation by about 7%, stating that despite the challenge of the activities, which is expected due to the lack of familiarity with physical programming, they find physics easy and would choose a physics-related job. Astalini, Kurniawan, Perdana, and Kurniawan (2019) link positive motivation to the importance of Physics to society, the scientific way of working in general, personal enjoyment, interest in increasing the time engaged in it, and interest in a career related to Physics subject.

It was also observed in the present research that both groups increased their time engaged in physics activities, with the experimental group reporting that they could perform exceptionally well on physics assignments and tests (in question

27, there was a 5% increase). The differences between the experimental and control groups can be attributed to the fact that the children in the experimental group begin to realize that they can achieve good performance in physics through physical programming. Also, success in using the materials can explain the reduction in stress levels on the one hand and the increase in self-confidence and internal motivation on the other, which was already mentioned (Love, 2023).

All children, regardless of group, improved their performance. Critical to using the proposed approach is confirming that project-based learning, as suggested by the learning sequence, improves learning (Kastner-Hauler et al., 2022). The difference observed in the post-test between the groups is slight (2.24%). A difference was found in the performance improvement rates: 27.04% for the control group and 30.28% for the experimental group.

In general, a significant effect of the use of the Micro:bit on the children's attitude was observed, even if this was not so strongly reflected in their performance. A finding confirms that various learning results arise, such as improving students' performance and technical skills, but also emotional results related to enhancing interest, motivation, and involvement (Ariza & Baez, 2021).

Also, the typical cognitive learning outcomes involve improving students' grades and technical skills. Regarding emotional learning outcomes, the increase in interest, motivation, and engagement is empirical evidence that needs more investigation (Kalogiannakis, Tzagkaraki, & Papadakis, 2021).

The students faced several challenges. Their interest in solving complex and frustrating problems was expected to reduce students' initial attitudes or motivation. On the contrary, the research showed that the students' attitudes and motivation remained high or slightly increased at the end of the activities. (Kaloti-Hallak, Armoni, & Ben-Ari, 2015).

Limitations – Perspectives

Regarding the use of the Micro:bit device, the main concern is related to the integrated sensors (and related concepts), which may remain black boxes for students of this age. This means there is not enough flexibility in expanding the variety and number of connected components to enhance the understanding of electricity. The attitude survey tool is designed to identify attitudes even if they change, but it does not explain why this might happen.

The results contribute to research dealing with the design of educational programs, emphasizing the pedagogy of Science and STEM courses and the use of educational robotics in the classroom. A combination of assessment methods, such as observation and interview, could capture deeper aspects of the subject being examined and information not easily discernible from the quantitative

approach.

Also, the investigation of this approach from the point of view of teachers would complement the related research with the rationale that if a teacher has sufficient expertise and self-confidence, he/she can open new paths for his/her students in the world of physical programming, even if not the ultimate coding expert.

References

Ariza, J. Á., & Baez, H. (2022). Understanding the role of single-board computers in engineering and computer science education: A systematic literature review. *ArXiv*. https://doi.org/10.1002/cae.22439

Astalini, Kurniawan, D. A., Perdana, R., & Kurniawan, W. (2019). Identification Attitudes of Learners on Physics Subjects. *Journal of EST (Educational Science And Technology)*, 5(1), 39-48. https://doi.org/10.26858/est.v5i1.8231

Chaldi, D., & Mantzanidou, G. (2021). Educational robotics and STEAM in early childhood education. *Advances in Mobile Learning Educational Research*, *1*(2), 72-81. https://doi.org/10.25082/AMLER.2021.02.003

Chiu, M.-C., Hwang, G.-J. & Tu Y.-F. (2022). Roles, applications, and research designs of robots in science education: a systematic review and bibliometric analysis of journal publications from 1996 to 2020. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2022.2129392

Foti, P. (2021). Exploring kindergarten teachers' views on STEAM education and educational robotics: Dilemmas, possibilities, limitations. *Advances in Mobile Learning Educational Research*, *1*(2), 82-95. https://doi.org/10.25082/AMLER.2021.02.004

Hillman, S.J., Zeeman, S.I., Tilburg, C.E. et al. (2016). My Attitudes Toward Science (MATS): the development of a multidimensional instrument measuring students' science attitudes. *Learning Environ Res*, 19, 203–219. https://doi.org/10.1007/s10984-016-9205-x

Kalogiannakis, M., & Papadakis, S. (2022). Educating preservice teachers to introduce educational robotics into their future preschool classroom. In *Proceedings of the 11th Virtual Edition of the International Conference New Perspectives in Science Education*, Italy-Florence.

Kaloti-Hallak, F., Armoni, M., & Ben-Ari, M. (2015). Students' attitudes and motivation during robotics activities. ACM *International Conference Proceeding Series*, 09-11-Nove, 102 – 110. https://doi.org/10.1145/2818314.2818317

Kastner-Hauler, O., Tengler, K., Sabitzer, B., & Lavicza, Z. (2022). Combined Effects of Block-Based Programming and Physical Computing on Primary Students' Computational Thinking Skills. *Frontiers in Psychology*, 13, 875382. https://doi.org/10.3389/fpsyg.2022.875382

Kvaššayová, N., Cápay, M., Petrík, Š., Bellayová, M., & Klimeková, E. (2022). Experience with Using BBC Micro:Bit and Perceived Professional Efficacy of Informatics Teachers. *Electronics,* 11(23), 3963.

Kyriazopoulos, I., Koutromanos, G., Voudouri, A., & Galani, A. (2021). Educational robotics in primary education: A systematic literature review. In St. Papadakis & M. Kalogiannakis (Eds.), *Handbook of research on using education robotics to facilitate student learning* (pp. 377–401). IGI Global. https://doi.org/10.4018/978-1-7998-6717-3.ch015

Love, T. S. (2023). Examining middle school students' attitudes toward computing after participating in a physical computing unit. *Interactive Learning Environments*, 1-20. https://doi.org/10.1080/10494820.2023.2194326

Méheut, M. & Psillos, D. (2004). Teaching–learning sequences: aims and tools for science education research, *International Journal of Science Education*, 26(5), 515–535. https://doi.org/10.1080/09500690310001614762

Min, S. H., & Kim, M. K. (2020). Developing Children's Computational Thinking through Physical Computing Lessons. *International Electronic Journal of Elementary Education*, 13(2), 183–198.

National Research Council. (2007). *Taking Science to School: Learning and Teaching Science in Grades K-8*. Washington, DC: The National Academies Press. https://doi.org/10.17226/11625

Petousi, V., & Sifaki, E. (2020). Contextualizing harm in the framework of research misconduct. Findings from discourse analysis of scientific publications, *International Journal of Sustainable Development, 23*(3/4), 149-174, https://dx.doi.org/10.1504/IJSD.2020.10037655

Przybylla, M., & Romeike, R. (2018, November). Impact of physical computing on learner motivation. In *Proceedings of the 18th Koli Calling International Conference on Computing Education Research* (pp. 1-10).

Salta, K., & Koulougliotis, D. (2015). Assessing motivation to learn chemistry: adaptation and validation of Science Motivation Questionnaire II with Greek secondary school students. *Chemistry Education Research and Practice*, 16, 237-250.

Sentance, S., Waite, J., MacLeod, E., & Yeomans, L. E. (2017). Teaching with physical computing devices: the BBC micro:bit initiative. *In Proceedings of 12th Workshop in Primary and Secondary Computing Education: WIPSCE '17*. https://doi.org/10.1145/3137065.3137083

Sheeran, P. (2002). Intention-behaviour relations: A conceptual and empirical review. *European Review of Social Psychology*, 12, 1-36.

Tai, R.- H., Ryoo, J. - H., Skeeles-Worley, A. et al. (2022). (Re-)Designing a measure of student's attitudes toward science: a longitudinal psychometric approach0. *IJ STEM Ed*, 9, (12). https://doi.org/10.1186/s40594-022-00332-4

Tzagaraki, E., Papadakis, S., & Kalogiannakis, M. (2022). Teachers' attitudes on the use of educational robotics in primary school. In *STEM, Robotics, Mobile Apps in Early Childhood and Primary Education: Technology to Promote Teaching and Learning* (pp. 257-283). Singapore: Springer Nature Singapore.

Tzimopoulos, N., Provelengios, P., Iosifidou, M. (2021). Implementation and evaluation of a remote seminar on the pedagogical use of educational robotics. *Advances in*

Mobile Learning Educational Research, *1*(2), 48-57. https://doi.org/10.25082/AMLER.2021.02.001

Vordou, E., & Romero, M. (2021). *Supporting Grammar Language Through a Micro:Bit Activity* (Report No. 0002, UCA - INSPE Académie de Nice). https://hal.archives-ouvertes.fr/hal-03408009v2

You, H. S., Kim, K., Black, K., & Min, K. W. (2018). Assessing Science Motivation for College Students: Validation of the Science Motivation Questionnaire II using the Rasch-Andrich Rating Scale Model. *Eurasia Journal of Mathematics, Science and Technology Education, 14*(4), 1161-1173. https://doi.org/10.29333/ejmste/81821

Stimulating Reflection Processes with an Eye-Gaze-Augmented Retrospective in Organic Chemistry

Axel Langner, Nicole Graulich

Justus-Liebig-University Giessen, Germany

Understanding complex representations and the respective implicit content they convey to solve domain-specific problems is a crucial skill in the professional development of an organic chemist. Although various studies indicate that students often struggle to derive implicit information from organic chemical representations, there is a lack of instructions to foster students' comprehension of such complex representations individually in the sense of differentiated instruction. A potential differentiated instructional approach may be reflecting on one's own problem-solving process, as critically questioning one's underlying assumptions to draw individual consequences for future problem-solving situations. Especially in domains with complex representations like Organic Chemistry, eye-tracking may support these reflection processes by providing insights into information inaccessible to a learner, such as the own gaze. Therefore, a qualitative, exploratory study has been conducted to investigate to what extent an eye-gaze-augmented retrospective can stimulate students to reflect on their problem-solving process. The findings suggest that the structural complexity of the observed reflection patterns does not per se imply productive, conceptual sensemaking.

Keywords: Eye-tracking, reflection, organic chemistry.

Introduction

Problem-solving in Organic Chemistry

Problem-solving can be seen as a 21st-century competence (Voogt & Roblin, 2012). Particularly in chemistry, problem-solving is linked to domain-specific representational competence (Kozma & Russell, 2005). In Organic Chemistry, symbolic-iconic representations containing explicit and implicit features of the chemical entities are used by chemists (Hoffmann & Laszlo, 1991) to reason about implicit properties and associated chemical reactivity to draw conclusions about the plausibility of chemical reaction mechanisms. This implicit information must be deduced from the explicit, visually decodable symbolic representation (Goodwin, 2010). However, learners in Organic Chemistry often struggle to derive implicit properties from organic chemical representations while solving problems, as they often rely on heuristics (Talanquer, 2014) or focus on single,

salient or superficial features (Graulich et al., 2019). Even though studies have shown that different learning instructions can support problem-solving skills of learners with a certain level of prior knowledge (*e.g.*, example-based learning as reviewed by van Gog and Rummel (2010)), there is a lack of differentiated instructional approaches in the sense of educational equity (cf. Dumont & Ready, 2023) that consider the needs of learners and support them in their individual problem-solving process.

Reflection as a differentiated instructional approach

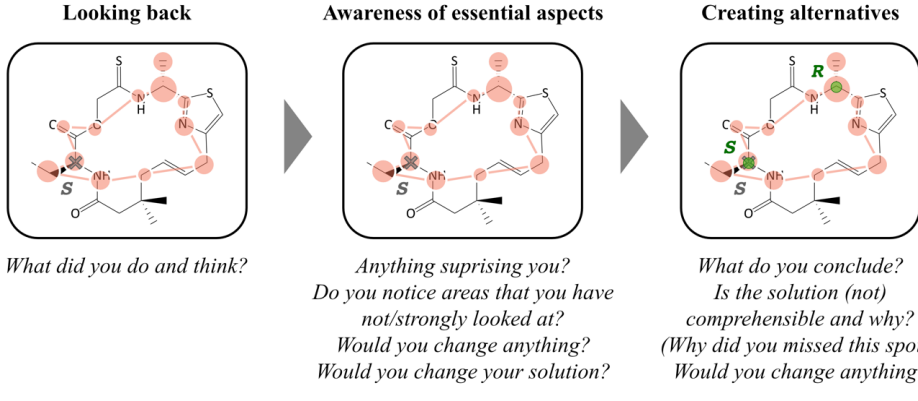
A potential differentiated instructional approach may be the reflection on one's own problem-solving process, as critically questioning one's underlying assumptions when applying professional knowledge can be a means to construct new knowledge (Ng et al., 2015). Thereby, reflection can be seen as actively and carefully weighing of beliefs or convictions by taking the underlying reasons of the reflection process and the subsequent conclusions into account (Dewey, 1933). In general, reflection can be simplified into the following phases: 1.) returning to an experience, 2.) evaluating the experience, and 3.) changing beliefs or behaviors (Atkins & Murphy, 1993). However, reflecting on your own may be challenging. Educators assisting reflection may promote the reflective learning by listening, guiding prompts, and reflective questions (Boud et al., 1985). Additionally, different reflection models, like the ALACT-model of Korthagen (1985), can be used to structure reflection processes. The ALACT-model comprises five phases: action (A), looking back on the action (L), awareness of essential aspects (A), creating alternative methods of action (C), and trial (T). Furthermore, reflection process may be accompanied by different tools. Besides traditional tools, such as journals or videos, sensor devices could capture and represent new information that usually would not be accessible to a learner (Fleck & Fitzpatrick, 2010). Especially in a domain with complex representations, like Organic Chemistry, eye-tracking technology could offer learners new information about their visual and conceptual processing. In an eye-gaze-augmented retrospective, learners could reflect on their problem-solving process, activate additional conceptual resources (diSessa, 1993; Hammer et al., 2005), carry out an productive sensemaking process (Odden & Russ, 2019), and ideally draw conclusions to inform future problem-solving situations.

An Eye-Gaze-Augmented Retrospective

In an eye-gaze-augmented retrospective learners are evaluating an experience retrospectively while watching an eye-gaze replay depicting their own eye-movements during this experience to reflect about it. By depicting conscious and unconscious fixations (*i.e.*, when the eyes fixate something) and saccades (*i.e.*, rapid shift between two fixations) in a dynamic manner (Holmqvist et al., 2011), the eye-gaze replay could offer learners new information about their

visual problem-solving actions (van Gog et al., 2005). Thereby, the eye-gaze-augmented retrospective could follow phases similar to the phases of the ALACT-model of Korthagen (1985). Originally used in teacher education, the phases of the ALACT-model are also applicable for problem-solving situations in science contexts as it focuses on essential aspects of an action to create new methods of action for future trials. Therefore, the phases of the ALACT-model were adapted for the eye-gaze-augmented retrospective: First, while watching their own eye-gaze-replay and response, students were prompted to describe what they did and thought during their problem-solving process (looking back). Secondly, students were asked reflective questions about their problem-solving process (awareness of essential aspects). Third and finally, while the task solution was presented, students were asked additional questions to compare their responses with the solution (creating alternatives) (see Figure 1).

Figure 1. Structure of the eye-gaze-augmented retrospective.



The retrospective was structured as a semi-structured interview, enabling the interviewer to address specific events and use generic (encourage to think aloud, *e.g.*, *Please describe as detailed as possible, what you did and thought.*) and direct prompts (contextualized prompts, *e.g.*, *What may be the reasons that you did not notice this spot?)* to encourage the reflection process (cf. Davis, 2003). As students may have a simplistic conception of reflection (common sense definition vs. academic definition) (Moon, 2004), students were not explicitly prompted to reflect. Instead, the given prompts and questions aimed to implicitly stimulate reflection. Additionally, guidelines and recommendations from Ericsson and Simon (1993) for retrospectives and from Fleck and Fitzpatrick (2010), Korthagen and Kessels (1999), and Sandars (2009) for reflection-based instructions were taken into account.

Research Aim and Design

The potential of such an eye-gaze-augmented retrospective to foster the comprehension of complex representations in the context of organic-chemical problem-solving tasks has not been investigated so far. Hence, we investigated in an exploratory study to what extent an eye-gaze-augmented retrospective stimulates students to reflect on their problem-solving process in Organic Chemistry. Hereby, we seek to answer the following research questions (RQ):

> RQ1: What patterns of reflection can be stimulated with an eye-gaze-augmented retrospective?
>
> RQ2: To what extent is the structural complexity of the reflection patterns associated with productive, conceptual sensemaking?

Therefore, twenty-two Organic Chemistry beginner students from a German university were prompted to solve three tasks about different contexts in Organic Chemistry (stereochemistry, Brønsted-acids, electrophilicity). While solving each task, students' eye-movements were recorded with an eye-tracker. Immediately after each task, the eye-gaze-augmented retrospective interview was carried out (see Figure 1).

Data Analysis

During the retrospective, students reflected on different key structural sites within the molecules. For each task, single reflection processes on two characteristic key structural sites were selected, resulting in a total of 132 analyzed reflection processes. The analysis of the reflection processes involved multiple steps addressing the research questions.

1. To characterize the reflection processes, each reflection process was coded dichotomously based on the presence or absence of characteristic reflection elements, using qualitative content analysis. The four codes used for the deductive qualitative content analysis were derived from the extensive literature on reflection: describing experiences (cf. Korthagen & Kessels, 1999; Nguyen et al., 2014), identifying issues (cf. Boyd & Fales, 1983; Schön, 1983), creating alternatives (cf. Fleck & Fitzpatrick, 2010; Korthagen, 1985), drawing consequences (cf. Atkins & Murphy, 1993; Moon, 2004)
2. Based on the combination of assigned codes, the reflection processes were categorized into emerging reflection patterns (RQ1).
3. Chosen examples were contrasted regarding structural complexity and productive, conceptual sensemaking (RQ2).

Summarized Results

RQ1: What reflection patterns can be stimulated with an eye-gaze-augmented retrospective?

Based on the occurrence of the four codes, five distinct reflection patterns with raising complexity were observed: descriptive reflection, non-elaborated reflection, alternative-centered reflection, consequence-centered reflection, elaborated reflection.

RQ2: To what extent is the structural complexity of the reflection patterns associated with productive, conceptual sensemaking?

By comparing reflection processes with high structural complexity, *i.e.*, an elaborated reflection, the results indicate that structural complex reflection patterns do not per se imply productive, conceptual sensemaking. For instance, in some elaborated reflection students could activate additional conceptual resources related to chemical concepts and derived the need to adjust their attention and application of chemical concept knowledge in the future. In contrast, in other elaborated reflection processes, students could not make sense of the task solution and concluded that they would just mark the same structure in the future, if they will recognize the same structure again. To summarize, despite that reflection processes had the same structural complexity, only some of the reflection processes were characterized by a productive, conceptual sensemaking process.

Conclusion

Through the eye-gaze-augmented retrospective, various reflection processes were stimulated, differing in their structural complexity and the extent of productive, conceptual sensemaking. The results of this exploratory study suggest that the eye-gaze-augmented retrospective demonstrates the potential of a differentiated instructional approach, but further optimizations are still needed. The use of scaffolding (Wood et al., 1976) for metacognitive support of one's reflection structure, or the use of machine learning feedback to adapt the retrospective to the learners' prior knowledge, could be the focus of future research.

Acknowledgement

We would like to thank the German Research Foundation DFG (Deutsche Forschungsgemeinschaft) for funding this research project (project number: 453926661). We also thank the students who voluntarily participated in the study, as well as everyone who was involved in this project.

References

Atkins, S., & Murphy, K. (1993). Reflection: a review of the literature. *Journal of Advanced Nursing*, *18*(8), 1188-1192. https://doi.org/10.1046/j.1365-2648.1993.18081188.x

Boud, D., Keogh, R., & Walker, D. (1985). *Reflection: Turning experience into learning*. Routledge.

Boyd, E. M., & Fales, A. W. (1983). Reflective Learning: Key to Learning from Experience. *Journal of Humanistic Psychology*, *23*(2), 99-117. https://doi.org/10.1177/0022167883232011

Davis, E. A. (2003). Prompting Middle School Science Students for Productive Reflection: Generic and Directed Prompts. *The Journal of the Learning Sciences*, *12*(1), 91-142. https://doi.org/10.1207/S15327809JLS1201_4

Dewey, J. (1933). *How we think*. Heath & Co Publishers.

diSessa, A. A. (1993). Toward an Epistemology of Physics. *Cognition and Instruction*, *10*(2/3), 105-225. https://doi.org/10.1080/07370008.1985.9649008

Dumont, H., & Ready, D. D. (2023). On the promise of personalized learning for educational equity. *npj Science of Learning*, *8*(1), 26. https://doi.org/10.1038/s41539-023-00174-x

Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis: Verbal Reports as Data*. The MIT Press.

Fleck, R., & Fitzpatrick, G. (2010). Reflecting on reflection: framing a design landscape. *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction* OZCHI '10, Brisbane, Australia.

Goodwin, W. (2010). How do Structural Formulas Embody the Theory of Organic Chemistry? *The British Journal for the Philosophy of Science*, *61*(3), 621-633. https://doi.org/10.1093/bjps/axp052

Graulich, N., Hedtrich, S., & Harzenetter, R. (2019). Explicit versus implicit similarity – exploring relational conceptual understanding in organic chemistry. *Chemistry Education Research and Practice*, *20*(4), 924-936. https://doi.org/10.1039/C9RP00054B

Hammer, D., Elby, A., Scherr, R. E., & Redish, E. F. (2005). Resources , framing , and transfer. In J. P. Mestre (Ed.), *Transfer of Learning from a Modern Multidisciplinary Perspective* (pp. 89-119). Information Age Publishing.

Hoffmann, R., & Laszlo, P. (1991). Representation in Chemistry. *Angewandte Chemie International Edition in English*, *30*(1), 1-16. https://doi.org/10.1002/anie.199100013

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.

Korthagen, F. A. (1985). Reflective teaching and preservice teacher education in the Netherlands. *Journal of teacher education*, *36*(5), 11-15. https://doi.

org/10.1177/002248718503600502

Korthagen, F. A., & Kessels, J. P. (1999). Linking theory and practice: Changing the pedagogy of teacher education. *Educational Researcher*, *28*(4), 4-17. https://doi.org/10.3102/0013189X028004004

Kozma, R., & Russell, J. (2005). Students Becoming Chemists: Developing Representationl Competence. In J. K. Gilbert (Ed.), *Visualization in Science Education* (pp. 121-145). Springer Netherlands. https://doi.org/10.1007/1-4020-3613-2_8

Moon, J. A. (2004). *A Handbook of Reflective and Experiential Learning: Theory and Practice*. Psychology Press.

Ng, S. L., Kinsella, E. A., Friesen, F., & Hodges, B. (2015). Reclaiming a theoretical orientation to reflection in medical education research: a critical narrative review. *Medical Education*, *49*(5), 461-475. https://doi.org/10.1111/medu.12680

Nguyen, Q. D., Fernandez, N., Karsenti, T., & Charlin, B. (2014). What is reflection? A conceptual analysis of major definitions and a proposal of a five-component model. *Medical Education*, *48*(12), 1176-1189. https://doi.org/10.1111/medu.12583

Odden, T. O. B., & Russ, R. S. (2019). Defining sensemaking: Bringing clarity to a fragmented theoretical construct. *Science Education*, *103*(1), 187-205. https://doi.org/10.1002/sce.21452

Sandars, J. (2009). The use of reflection in medical education: AMEE Guide No. 44. *Medical Teacher*, *31*(8), 685-695. https://doi.org/10.1080/01421590903050374

Schön, D. A. (1983). *The Reflective Practitioner*. Temple Smith.

Talanquer, V. (2014). Chemistry Education: Ten Heuristics To Tame. *Journal of Chemical Education*, *91*(8), 1091-1097. https://doi.org/10.1021/ed4008765

van Gog, T., Paas, F., Van Merriënboer, J. J., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied*, *11*(4), 237. https://doi.org/10.1037/1076-898X.11.4.237

van Gog, T., & Rummel, N. (2010). Example-Based Learning: Integrating Cognitive and Social-Cognitive Research Perspectives. *Educational Psychology Review*, *22*(2), 155-174. https://doi.org/10.1007/s10648-010-9134-7

Voogt, J., & Roblin, N. P. (2012). A comparative analysis of international frameworks for 21st century competences: Implications for national curriculum policies. *Journal of Curriculum Studies*, *44*(3), 299-321. https://doi.org/10.1080/00220272.2012.668938

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of child psychology and psychiatry*, *17*(2), 89-100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

The Use of Augmented Reality in Biology Subjects in the Context of Technology Integration into Education

Sonay Şahin Aydemir[1], Esra Kızılay[1], Aslı Saylan Kırmızıgü[1], Nagihan Tanık Önal[2]

[1]Erciyes University, Türkiye

[2]Niğde Ömer Halisdemir University, Türkiye

Technology integration into education has long been on educators' agenda and is intriguing for teachers and students. However, advancements in technology have also changed the technologies used in education. One such new technology is Augmented Reality (AR), which allows students to learn without disconnecting from the real world. AR applications are practical learning tools when attention is paid to their pedagogical dimension. Teachers play a crucial role in achieving positive outcomes. Therefore, exploring teachers' views on the use of AR is considered necessary. Additionally, AR has many advantages in teaching abstract and not easily observable topics in a regular classroom setting, making it particularly suitable for use in biology education. However, research in this area has primarily focused on physics topics, with limited studies addressing biology subjects. In this context, the current research aims to examine science teachers' views on using AR in teaching biology subjects in detail. Based on the phenomenological design, the qualitative research method involved six science teachers working in Kayseri, Türkiye. Data for the research were collected through semi-structured interviews conducted with the participants on an online platform. Descriptive and content analysis were then used together to analyze the data. The research findings indicate that participants generally have positive views on integrating technology and using AR in teaching biology subjects. Participants believe that using AR in biological issues has advantages such as making abstract concepts concrete, being fun and engaging, providing lasting learning in students, and increasing motivation for the course. However, there are also opinions suggesting disadvantages, such as the continuous use leading to routine and eliminating the novelty effect and the requirement of specific technological knowledge of usage. According to the findings, teachers mostly use AR applications in biology subjects related to "Systems," "Cell and Organelles," and "Diversity of Living Organisms." The preferred AR applications are Quiver, Anatomy 4D, Animal 4D+, and Humanoid 4D+. Based on the research results, recommendations include the widespread adoption of AR in teaching biology subjects and providing practical training to teachers in this regard.

Keywords: Technology integration in education, augmented reality, biology subjects.

Introduction

Integrating technology into education is highly intriguing for both teachers and students, but it also brings specific challenges. Overall, technology integration in education supports teachers in creating a more constructivist learning environment, facilitating increased student engagement, better learning outcomes, and ultimately enhancing the quality of instruction (Linardatos & Apostolou, 2023). Teachers have been incorporating technology into their instruction for an extended period to improve the learning environment. However, with technological advancements, the tools used in education have evolved and become digitized, similar to how versatile whiteboards have replaced traditional chalkboards (Farjon et al., 2019). One of the current technologies applicable in education is Augmented Reality (AR). AR is a system based on the principle of combining computer-generated objects with the natural world (Akçayır & Akçayır, 2017). With its features, AR attempts to overcome the limitations of environments where students have no interaction with the real world, similar to virtual reality (İbili, 2013). When examining relevant studies, it is observed that AR applications in education have positive effects on student interest and motivation in the course (Delello, 2014), facilitate the teaching of abstract concepts (Cai et al., 2014), and make it possible and practical to teach concepts that are not visible in the real world or cannot be experienced in a classroom setting (Abdüsselam, 2014). According to Siegle (2019), AR is a new and exciting learning tool for students. Considering all these positive outcomes, using AR in education is recommended (Fleck & Simon, 2013).

AR materials stand out within the field of education, especially in the science education domain, which includes numerous abstract concepts that students often find challenging (Karal & Abdüsselam, 2015; Tanık Önal et al., 2021). However, these outcomes are valid for the successful implementation of technology integration. Teachers' frequent use of information and communication technologies does not necessarily mean that technology integration is successful. The number of teachers who can achieve this is relatively low (Farjon et al., 2019). Mishra and Koehler (2006) proposed a theoretical framework called Technological Pedagogical Content Knowledge (TPCK) for teachers to achieve full technology integration. This model emphasizes the importance of developing technical skills, pedagogical practices, and content knowledge in pre-service teachers and integrating these three components relaxed and harmoniously (Koehler & Mishra, 2009). In this context, for teachers to effectively use AR in technology integration in education, it is not enough for them to know this method or be willing to use it (So & Kim, 2009). Technology integration is

a complex process, and factors such as teachers' technological pedagogical content knowledge, attitudes, competencies, experiences, and belief systems play a pivotal role in their successful implementation. Therefore, a thorough examination of teachers' perspectives on the subject is deemed necessary. Additionally, the fact that AR is a current technology for science education practices the limited number of studies focused on AR and biology subjects (Arıcı, 2021; Gnidovec et al., 2020), and the teachers being implementers of educational programs position this research as unique and significant. With this perspective, this research aims to thoroughly examine science teachers' views on using AR in teaching biology subjects using a qualitative method.

Method

Research Design

This research was conducted using a qualitative research method based on the phenomenological design. Phenomenology aims to deeply examine the meaning and nature of individuals' experiences (Patton, 2014). This design was chosen since the study aims to examine the views of science teachers who have a specific experience with AR usage in their classes regarding the use of AR in Biology subjects.

Participants

The participants in the research were determined using purposeful sampling method, precisely the criterion sampling method. The criteria for selection were being a science teacher who uses AR applications in their classes and being willing to participate in the research. As a result, six science teachers from middle schools in the central and district areas of Kayseri province voluntarily participated in the research in the 2021-2022 academic year. Some demographic characteristics of the participants are provided in Table 1.

Table 1. Demographic characteristics of the participants

Participants	Gender	Years of work experience	Code names
1st participant	Female	9	Ö1
2nd participant	Female	9	Ö2
3rd participant	Female	8	Ö3
4th participant	Male	7	Ö4
5th participant	Male	10	Ö5
6th participant	Male	13	Ö6

Data Collection Tool

The data for the research were collected through a semi-structured interview form

between April and May 2022. Before the interviews, the form was developed by the researchers by reviewing the relevant literature in line with the research questions. The draft form was presented to three experts with a Ph.D. in science education for their opinions. Based on the expert opinions, some items were updated in terms of language and expression features, and the final version of the form was determined. The structure consists of three main sections and several sub-sections.

Data Collection

The interviews were conducted online using the Zoom program, and they were recorded to prevent data loss. Prior permission was obtained from the participants to record the interviews. Participants were informed that they could take a break or end the interview on time. This approach aimed to minimize any potential discomfort the interviews might cause participants. Each interview had an average duration of 25 minutes. Throughout the interviews, efforts were made to ensure that the researcher did not influence the participants' responses. The researcher listened without intervening in the answers, avoiding any affirmative or negative attitudes and behaviors.

Data Analysis

Before analyzing the research data, the transcriptions of the interviews were carefully converted into MS Word documents. Subsequently, each document obtained from every interview was assigned code names from Ö1 to Ö6. For data analysis, Creswell's (2008) content analysis approach was employed. In this regard, the coders established a template for organizing the data after thoroughly reading all the transcripts twice. Then, each coder independently conducted the coding process. After coding, similar codes were grouped together. Once the coding process was completed, themes were developed.

Findings

Teachers' Views on the Integration of Technology into Science Education

This theme consists of the participants' views on integrating technology into science education, divided into two sub-themes: Advantages and disadvantages. Participants' opinions in the advantages sub-theme include:

- the positive effects of technology integration on the learning process,
- alignment with the characteristics of contemporary students and
- the nature of science education.

According to science teachers, the integration of technology into their classes has advantages such as ensuring lasting learning (f=5), making abstract concepts

concrete (f=3), being attention-grabbing (f=2), facilitating active participation (f=2), enrichment (f=2), and being fun (f=2). Below are some quotes illustrating this sub-theme:

Ö1: *"I find it positive in terms of making the lesson more concrete. I believe it makes students more active in the class and facilitates the processing of topics."*

Ö2: *"Thanks to technology, we can access information faster and easier. This adds richness to the learning processes, providing more meaningful learning both from a teaching and learning perspective."*

Ö6: *"It enriches the content of the lesson."*

Additionally, four teachers expressed that science education is highly suitable for technology integration, and two teachers mentioned the positive reflections of technology integration in science classes parallel to the characteristics of today's children. An example quote encompassing both views is as follows:

Ö5: *"As science teachers, maybe we are the luckiest regarding technology. Because our subjects are parts of life, it is crucial to integrate technology. The new generation, the Z generation, is very inclined towards technology. They are constantly looking for technological content in the class, and seeing it affects them even more."*

The disadvantages sub-theme regarding the technology integration is categorized into application-related disadvantages and teacher-related disadvantages. Application-related disadvantages include cost (f=3), application difficulty (f=2), security issues (f=2), and overuse (f=2). Teacher-related disadvantages consist of views on teachers' lack of knowledge regarding technology applications. Some example quotes representing these views are:

Ö1: *"There can be difficulty in application, especially in crowded classes."*

Ö2: *"I think there may be problems regarding cost."*

Ö3: *"It should be done under parental supervision."*

Ö6: *"Overuse can be a disadvantage."*

Furthermore, within this theme, teachers were asked if they use digital applications in their classes, to which all participants answered yes. When asked about the most frequently used technologies, Interactive Boards, in line with the recommendations of the Turkish Ministry of National Education, were the most commonly mentioned technology (f=4).

Teachers' Views on the Use of Augmented Reality in Biology Topics

This theme focuses on teachers' opinions regarding the use of AR, specifically in

biology topics within science classes. Initially, all participants expressed positive views on using AR in biology topics. The responses, similar to the theme of technology integration, can be categorized into two sub-themes: advantages and disadvantages. According to science teachers, the use of AR in biology topics has advantages such as making abstract concepts concrete (f=4), being fun (f=4), ensuring lasting learning (f=3), providing motivation (f=3), and being attention-grabbing (f=2). Below are example quotes representing these views:

Ö4: *"It makes the subject more concrete."*

Ö1: *"It makes it both fun and educational."*

Ö2: *"It makes learning lasting."*

Ö5: *"It arouses curiosity and motivates."*

Ö6: *"It is very attention-grabbing."*

When examining the participants' views on the disadvantages of using AR in biology topics, it is observed that the responses are generally shaped in the context of technology usage, and there is no direct emphasis on a disadvantage specific to the feature of AR. The opinions include drawbacks such as the continuous use leading to monotony (f=4), requiring specific technological knowledge (f=3), and design differences (f=2). Below are example quotes illustrating these views:

Ö1: *"Continuous use can make it boring and mundane."*

Ö2: *"Applications may require expertise."*

Ö5: *"Possible differences between reality and visuals in the application's content."*

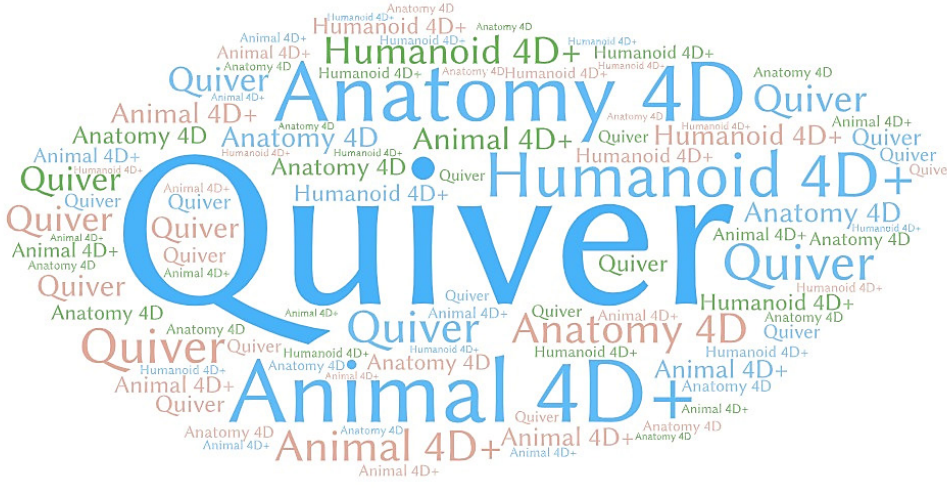
Next, an attempt was made to determine which biology topics the participants used AR. In response to this question, the answers were mainly related to the Systems topic (f=6), followed by Cell and Organelles (f=3), and Diversity of Living Organisms (f=2). The following quote exemplifies these views:

Ö1: *"I think it suits the Systems topic, Diversity of Living Organisms, and Cell and Organelles. The Systems topic allows for the detailed examination of all organs in a clean and hygienic environment. Diversity of Living Organisms enables children who have never seen different living organisms to see and experience different species in a three-dimensional, close-to-real way. Cell and Organelles allow schools without microscopes to easily examine the cell and organelles. In schools with microscopes, it supports the topic, making it more effective and enjoyable."*

Finally, the science teachers were asked about the AR applications they most commonly used in teaching biology topics. The most preferred AR applications

for teaching biology topics were Quiver: Plant and Animal Cell (f=5), Anatomy 4D (f=4), Animal 4D+: Animals (f=3), and Humanoid 4D+: Systems in our Body (f=2). Figure 1 illustrates these findings.

Figure 1. Most preferred AR applications in biology



Discussion, Conclusion and Recommendations

This qualitative research aims to determine science teachers' views on the use of Augmented Reality (AR) in teaching biology topics within the context of integrating technology into education. The initial finding from the study indicates that teachers generally have a positive outlook on integrating technology into education. Teachers believe that Science classes are highly suitable for technology integration, aligning with the characteristics of the new generation. Participants believe that integrating technology into education can contribute to cognitive (such as lasting learning and concretization) and affective (such as attention-grabbing, fun) domains in students. Similarly, Zovko (2016) notes that technology facilitates learning and teaching. Additionally, participants mention various disadvantages of technology integration, including cost, implementation difficulty, and teachers' need for knowledge in technology usage. The participants' perspectives are likely to influence their technology integration into their own classes, as technology integration is reported to be influenced by cognitive and affective factors (Yıldız Durak, 2019).

As for the focus of the current research on the use of AR in biology topics, participants expressed positive views. Furthermore, all teachers reported using AR applications to teach some biology topics. Upon closer analysis, teachers'

AR practices are considered to be generally at the micro-level, which represents the first level of integrating technology (Wang & Woo, 2007). In other words, teachers commonly use technology, specifically AR, in certain in-class activities to help students better understand some concepts.

Teachers believe that using AR in biology topics concretizes abstract concepts, is fun and engaging, ensures lasting learning in students, and enhances students' motivation for the subject. Similarly, in the literature, AR applications are noted to support both cognitive and affective learning domains (Ibáñez et al., 2014). The current study's findings align with other studies in the literature (Di Serio et al., 2017; Hwang et al., 2016; Radu, 2014). However, participants also mention disadvantages, such as the continuous use leading to monotony and the requirement for specific technological knowledge. It is noteworthy that, unlike the general literature, participants did not emphasize infrastructure and hardware deficiencies, such as the lack of Turkish language support in some AR applications or the absence of a mobile device for every student.

Another result obtained from the current research indicates that teachers extensively use AR applications in the biology topics of "Systems," "Cells and Organelles," and "Diversity of Living Organisms." In the literature, a study by Gnidovec et al. (2020) developed an AR application for teaching the circulatory system, proving its effectiveness. The features of AR, such as representing objects too small to be observed, showing situations that are not observable in a regular classroom, and concretizing abstract concepts, highlight the importance of AR in teaching these topics (Arıcı, 2021).

The most preferred AR applications for teaching biology topics among teachers are Quiver: Plant and Animal Cell, Anatomy 4D, Animal 4D+, and Humanoid 4D+: Systems in our Body. The easy accessibility, compatibility with different operating systems, and popularity of these applications may have contributed to their increased usage. For instance, Quiver is mentioned as a free, widely compatible with operating systems, easy-to-use, learning-facilitating, and highly interactive AR application (Karadayı Taşkıran et al., 2015).

Based on the current research results, it is recommended to promote the use of AR further in teaching biology topics, with scientific studies testing the effectiveness of AR usage in different class levels and various topics. Scientific studies should evaluate the effectiveness of different AR applications using robust methodologies that combine different data collection tools. Finally, providing teachers with practical training in technology integration is also recommended.

References

Abdüsselam, S. M. (2014). Teachers' and students' views on using augmented reality environments in physics education: 11th grade magnetism topic example. *Pegem Journal of Education & Instruction, 4*(1), 59-74.

Akçayır, M., & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review, 20*, 1-11.

Arıcı, F. (2021). *An Investigation of the effectiveness of problem based learning method enhanced with augmented reality in the teaching of cell and divisions subject* [Unpublished doctoral dissertation]. Ataturk University, Erzurum.

Cai, S., Wang, X., & Chiang, F. K. (2014). A case study of augmented reality simulation system application in a chemistry course. *Computers in Human Behavior, 37*, 31-40. https://doi.org/10.1016/j.chb.2014.04.018

Cresswell, J. W. (2008). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (Third edition). New Jersey: Pearson Education Ltd.

Delello, J. A. (2014). Insights from pre-service teachers using science-based augmented reality. *Journal of Computers in Education, 1*(4), 295-311.

Di Serio, A., Ibanez, M. B., & Kloos, C. D. (2013). Impact of an augmented reality system on students' motivation for a visual art course. *Computers & Education, 68*, 586-596.

Farjon, D., Smits, A., & Voogt, J. (2019). Technology integration of pre-service teachers explained by attitudes and beliefs, competency, access, and experience. *Computers & Education, 130*, 81-93.

Fleck, S., & Simon, G. (2013). *An augmented reality environment for astronomy learning in elementary grades: An exploratory study.* In Proceedings of the 25th Conference on l'Interaction Homme-Machine (pp. 14-22).

Gnidovec, T., Žemlja, M., Dolenec, A., & Torkar, G. (2020). Using augmented reality and the structure–behavior–function model to teach lower secondary school students about the human circulatory system. *Journal of Science Education and Technology, 29*(6), 774-784.

Hwang, G-J., Wu, P.-H., Chen, C.-C., & Tu, N.-T. (2016). Effects of an augmented realitybased educational game on students' learning achievements and attitudes in real-world observations. *Interactive Learning Environments, 24*(8), 1895-1906. https://doi.org/10.1080/10494820.2015.1057747

Ibáñez, M. B., Di-Serio, Á., Villarán-Molina, D., & Delgado-Kloos, C. (2014). Augmented reality-based simulators as discovery learning tools: An empirical study. *IEEE Transactions on Education, 58*(3), 208-213.

Ibili, E. (2013). *Development, implementation and assessment of the effect augmented reality on geometry teaching materials for geometry classes* [Unpublished doctoral dissertation]. Gazi University, Ankara.

Karadayı Taşkıran, A., Koral Gümüşoğlu, E., & Bozkurt, A. (2015). *The use of augmented reality application in foreign language teaching.* Academic Informatics Congress,

Eskişehir.

Karal, H., & Abdüsselam, M. S. (2015). Augmented reality. In B. Akkoyunlu, A. İşman, & F. Odabaşı (Eds.), *Educational Technology Readings 2015* (p. 149-176). Ayrıntı Press.

Koehler, M., & Mishra, P. (2009). What is technological pedagogical content knowledge (TPACK)?. *Contemporary Issues in Technology and Teacher Education, 9*(1), 60-70.

Linardatos, G., & Apostolou, D. (2023). Investigating high school students' perception about digital comics creation in the classroom. *Education and Information Technologies, 28,* 10079-10101. https://doi.org/10.1007/s10639-023-11581-3

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record, 108*(6), 1017-1054.

Patton, M. Q. (2014). *Qualitative research & evaluation methods: Integrating theory and practice.* Sage publications.

Radu, I. (2014). Augmented reality in education: A meta-review and cross-media analysis. *Personal and Ubiquitous Computing, 18*, 1533-1543.

Siegle, D. (2019). Seeing is believing using virtual and augmented reality to enhance student learning. *Gifted Child Today, 42*(1), 46-52. https://doi.org/10.1177/1076217518804854.

So, H.-J., & Kim, B. (2009). Learning about problem based learning: Student teachers integrating technology, pedagogy and content knowledge. *Australasian Journal of Educational Technology, 25*(1), 101-116.

Tanık Önal, N., & Önal, N. (2021). The effect of augmented reality on the astronomy achievement and interest level of gifted students. *Education and Information Technologies, 26*(4), 4573-4599.

Wang, Q., & Woo, H. L. (2007). Systematic planning for ICT integration in topic learning. *Journal of Educational Technology & Society, 10*(1), 148-156.

Yıldız Durak, H. (2019). Modeling of relations between K-12 teachers' tpack levels and their technology integration self-efficacy, technology literacy levels, attitudes toward technology and usage objectives of social networks. *Interactive Learning Environments, 29*(7), 1136-1162.

Zovko, V. (2016). ICT-enabled education need for paradigm shift. *Croatian Journal of Education, 18*(2), 145-155. doi: 10.15516/cje.v18i0.2175

Added Value of Google Lens For Identification of Marine Organisms

Vida Lang[1], Andrej Šorgo[1,2]

[1]Faculty of Natural Sciences and Mathematics, University of Maribor, Slovenia

[2]Faculty of Electrical Engineering and Computer Science, University of Maribor, Slovenia

Biodiversity can be recognized as a natural and cultural heritage that must be preserved. In field work with organisms, everything begins with their identification, and traditionally identification keys have been used in dichotomous and pictorial form. With the advent of mobile devices, new forms of identification based on image recognition and artificial intelligence have emerged. The 83 students were given the task of identifying marine organisms using the book, the Internet, and the smartphone application Google Lens. After completing the task, the differences between the three methods were evaluated. We did not find any significant differences in the quality of species identification, so all methods can be used interchangeably. Based on their experiences with all methods, students determined that the least difficult method for identifying marine organisms was using the book and the most difficult method for identifying marine organisms was using the Google Lens application. It was found that students preferred the Google Lens application even though they found it to be the most difficult method. The added value of the Google Lens application is that it is always accessible on smartphones. The students expressed great satisfaction with the experience of the workshop.

Keywords: Google Lens, marine organisms, smartphones and tablets.

Introduction

In the field of biology education, the concept of biodiversity is a cornerstone, although its explicit nomenclature is not always found in official documents. Biodiversity is multifaceted and is recognized at three basic levels, as Janžekovič (2022) has pointed out: the genetic level, the species level, and the community level. In the educational landscape, biodiversity in all its complexity is more than just a scientific topic or an explanation of natural heritage. It is also an integral part of our cultural heritage. This assertion is based on the understanding that the natural environment and human creations together form a common heritage-a community heritage rich in both tangible and intangible dimensions (Ahmad, 2006; UNESCO, 2023).

According to the definition of UNESCO (2023a), cultural heritage is the heritage that is passed down from the past, experienced in the present, and entrusted to future generations. Within this definition, the field of biological (science) education becomes a vessel for transmitting knowledge about natural heritage, often referred to as "nature's most precious gifts to humanity" (UNESCO, 2023b). Although cultural heritage is not always explicitly mentioned in official biology curricula, it is important to recognize that cultural heritage and the environment are inextricably linked. Therefore, the teaching of biology (science) should sit alongside other school subjects such as visual arts, music, mother tongue, history, and geography, where topics related to cultural heritage have traditionally been part of the curriculum. It is argued that many natural entities and ecosystems - from individual organisms and plant species to animal breeds, habitats, and even landscapes - have been shaped by human influence. Without human intervention, they are at risk of extinction. Protecting these organisms, habitats, landscapes, and practices therefore requires active involvement and regulation of human activities. Otherwise, the legacy of the past may be destruction and degradation. To ensure their preservation and distribution in altered landscapes, education must incorporate both tangible entities and intangible practices (Lowenthal, 2005). As Heberlein (2012) notes, those seeking to protect natural heritage must simultaneously seek solutions at three levels: technical, structural, and cognitive. The importance of these levels, which emphasize knowledge, skills, and attitudes for recognizing and conserving the diversity of the natural and human-made environment, has been highlighted. This is particularly evident in the conservation of extensive grasslands, one of the most threatened ecosystems in Europe, where the abandonment of traditional practices is a threat and considered part of the cultural heritage.

For students to acquire knowledge and skills that go beyond the classroom, teaching lifelong skills to conserve or enhance biodiversity is essential. Instruction must include familiarity with organisms, direct interaction with organisms, and practices relevant to those organisms.

In the context of field work with organisms, the journey begins with their identification. Traditionally, identification keys in various forms have served this purpose (Anđić et al., 2019). However, with the advent of mobile devices, new identification methods have emerged that rely on image recognition and artificial intelligence. Building on the experience gained with applications such as PlantNet (Lang & Šorgo, 2022), the goal of this study is to determine the preferred form of marine organism identification among secondary school students. By exploring their inclinations, we aim to advance understanding of how best to engage the next generation in biodiversity exploration and conservation.

Research Method and Design

During the 10-day Red Cross summer camp on the Adriatic Sea, a workshop was organized as part of the recreational activities for school children. A total of 83 schoolchildren, 55.4% of whom were boys and 44.6% girls, aged 11 to 16, voluntarily participated in this workshop, which was aimed at exploring the diversity of plants and animals in the coastal region.

The workshop lasted 1.5 hours and was repeated 10 times to accommodate all students. There were 9 students in each workshop session, and they were divided into 3 subgroups. Each subgroup was given specific instructions, which consisted of using a net to catch at least six different marine organisms and then identifying them using three different methods:

a) Image identification key.

b) Smart cell phone or tablet with internet access.

c) Smart cell phone or tablet with the Google Lens application.

For each method, there was a strict time limit of 15 minutes. After the allotted time, the subgroups switched to ensure that all students had the opportunity to use all three identification methods.

The tablets and smartphones required for this exercise were provided by the instructors. Each student was given a worksheet with step-by-step instructions that guided them through the identification process. After identification, the captured marine organisms were released back into their natural habitat. Students were then tasked with completing the opinion section of the provided instrument.

The correctness of the marine organism identification was rated on a scale of 0 to 3 using the following criteria:

a) Rank 0: Awarded to students who completely misnamed the organism or did not provide the answer.

b) Rank 1: Awarded to students who correctly identified at least the group to which the marine organisms belong or provided one of the correct names, regardless of whether it is Slovenian, professional, or English.

c) Rank 2: Reserved for students who correctly identified the organisms to the respective family or correctly gave at least two of the names (Slovenian, professional or English).

d) Rank 3: Reserved for students who correctly determined the name and species of the organism and correctly spelled all three names.

The goal of this workshop was to provide students with a hands-on experience in exploring and identifying the rich biodiversity of coastal plant and animal life,

while evaluating the effectiveness of various identification methods in enhancing their learning experience.

Statistical Procedure

In this study, different statistical methods were used to answer the research questions. Data analysis was performed using the statistical software IBM SPSS, version 28.0. The following is a summary of the main statistical procedures:

Descriptive statistics: Basic descriptive statistics were used to examine differences between participants' prior knowledge and their success in identifying marine organisms. These likely included measures such as means, standard deviations, and possibly inferential statistics such as t-tests or ANOVA to determine statistical significance.

Reliability Analysis (McDonald's Delta): To assess the reliability of the measurement scale used in the study, McDonald's delta was calculated. A satisfactory delta value indicates the internal consistency of the scale. Even though an increase in alpha was expected, no items were removed from the item pool to maintain the width of the scale. This indicates that the items in the scale were considered reliable and consistent for measuring the constructs of interest.

Principal Component Analysis (PCA): PCA analysis was conducted to reduce the data and identify the underlying components or factors. In this case, two components were apparently extracted that together explained 69.1% of the variance in the data. This analysis likely helped to simplify the data structure and reveal patterns or relationships among the variables.

Effect size (Cohen's h): Cohen's h was calculated to assess the relationship between groups. Effect sizes provide valuable information about the practical significance or magnitude of observed differences between groups. This measure helps researchers and readers understand the strength of the relationships or effects found in the study.

Statistical procedures were essential for conducting a rigorous and comprehensive analysis of the data collected in the study. They allowed researchers to draw meaningful conclusions and findings regarding participants' prior knowledge, identification success, scale reliability, data structure, and the impact of the Google Lens smartphone application on participants' satisfaction during the workshop.

Results

Successful identification of marine organisms

The table 1 shows how well the three identification methods (book, Internet,

Google Lens) helped participants correctly identify specific marine organisms. It provides information on the effectiveness of each method and highlights improvements compared to participants' prior knowledge, as indicated in the "Diff" column. This information is important for evaluating the educational impact of the workshop and the usefulness of the various identification methods.

Table 1. Frequency (%) of marine organisms correctly identified using three identification methods (N= 83).

Methods	Marine organism (MO)	Previous identification [%]	Professional name [%]	Slovenian name [%]	English name [%]	Diff*
Exercise 1: Book	MO1	79.5	77.1	86.7	15.7	7.2
Exercise 2: Internet	MO2	63.9	72.3	75.9	61.4	12.0
Exercise 3: Google Lens	MO3	73.5	71.1	80.7	67.5	7.2

*(Note: *Difference between the highest results of identification and Previous recognition.)*

Table 1 provides an overview of the success rates associated with three different exercise methods for marine organism identification. It is noteworthy that a significant proportion of students (over 63%) were shown to be able to identify marine organisms caught in a net without consulting identification books. However, when they used the identification key from the book, the success rate was highest when they correctly identified the Slovenian name of the marine organism. Conversely, the table shows that it was more difficult for the students to correctly identify the English name. This is partly because many Slovenian identification books lack a reference to the English name.

To assess progress, we calculated the difference between the best identification result and the students' prior knowledge of the organism's name. A larger difference indicates greater progress. The results suggest that the use of the free Internet as a resource contributed to a modest improvement in students' marine organism identification skills, although this progress was not particularly large.

The evaluation of the correctness of the plant identification was based on a point system with four levels from 0 to 3. Here is a breakdown of the evaluation criteria:

Rank 0: Awarded when a student named the plant organism completely incorrectly or left the answer blank.

Rank 1: Awarded to answers where the student correctly stated at least the group to which the plant belongs or one of the names (Slovenian, technical, or English).

Rank 2: Awarded to students who correctly assigned the organisms to their family or correctly stated at least two of the names (Slovenian, professional, or English).

Rank 3: Awarded to students who correctly identified the name and species of the plant organism and correctly spelled all three names.

Table 2 shows the results of this evaluation and the distribution of responses among the different ranks. The results show that marine organisms were identified with a high degree of accuracy using all three identification methods. Specifically:

- When using the book, 71.1% of the students identified the marine organisms completely correctly.
- When using the Internet method, 69.9% of the students succeeded in identifying the organisms completely and correctly.
- The Google Lens application also proved effective, with 71.1% of all students able to name the marine organisms completely and correctly.

Results demonstrate the overall competency of the students in plant identification and suggest that the three different identification methods provided comparable results in terms of correct identifications.

Table 2. Ranked the correctness (%) of a given name in 0 – 3 ranks (N= 83).

	Exercise 1: Book [%]	Exercise 2: Internet [%]	Exercise 3: Google Lens [%]
Rank 0	9.6	16.9	8.4
Rank 1	8.4	3.6	12.0
Rank 2	10.8	9.6	8.4
Rank 3	**71.1**	**69.9**	**71.1**

(Note: rank 0 = Wrong identification or blank answer, rank 3 = Complete identification of marine organism.)

The results of the first section of the study show that there are no significant differences in the identification of marine organisms when different methods are used. The percentage of organisms correctly named in their entirety is remarkably constant for all three methods, with the following results: Book method: 71.1%, Internet method: 69.9% and Google Lens method: 71.1%.

These results indicate that there is little difference in the accuracy of marine

organism identification between the three methods. Students achieved similar levels of success with each of these methods, highlighting their effectiveness in facilitating correct identifications.

Opinions about the use of smartphones and tablets in the worksphop

The first question of the questionnaire was designed to determine students' attitudes and opinions regarding the use of smartphones or tablets during the workshop. From the data presented in Table 3, the students have a positive attitude towards the use of smartphones during the fieldwork. They perceive the use of smartphones as effective and understandable. In addition, they express the view that the inclusion of smartphones in the exercise was enjoyable, instructive, and enlightening. These data suggest that students generally have a positive attitude toward integrating technology, particularly smartphones and tablets, into their learning experiences.

Table 3. Descriptive analysis of the question of satisfaction with the use of a smartphone for marine organism identification in the workshop (N = 83).

Due to the use of the tablet to identify marine organisms, the exercise was:	Mean	Med	Mode	SD
Succesfull	6.3	7	7	1.00
Understandable	6.2	6	7	1.03
Fun	6.1	6	7	1.24
Educational	6.1	6	7	1.07
Instructive	6.0	6	6	1.24

(Note. 1- do not agree at all, 2- disagree, 3- partially disagree, 4- neither agree nor agree, 5- partially agree, 6- agree, 7- completely agree.)

Perceived difficulty with the method

The second part of the questionnaire was designed to determine which method of identifying marine organisms used during the workshop students were most satisfied with (see Table 4).

In response to the question about which method they had the least difficulty with, it's noteworthy that 41.0% of the respondents indicated that they had the least difficulty when identifying marine organisms using the book. When asked about the most problematic method for identifying marine organisms, Google Lens proved to be the method that caused the most problems with 47.0% of the participants.

When asked about their preferences for future identification methods, 41.0% of

respondents expressed a desire to use the Google Lens application, while 33.7% indicated a desire to rely on traditional books.

Table 4. Analysis of the problem of difficulties in the implementation of plant identification with three different methods (N= 83).

Questions	Exercise 1: Book [%]	Exercise 2: Internet [%]	Exercise 3: Google Lens [%]	Mode
With which exercise did you have the least problems?	41.0	25.3	33.7	1
Which exercise did you have the most difficulty with?	32.5	20.5	47.0	3
If you had to do the exercise again, which execution method would you choose?	33.7	25.3	41.0	3

Table 5: Ranked the marine organism identification methods from least difficult (1) to most difficult (3).

Method of performing the exercise	1	2	3	Mode
Exercise 1: Book	43.3	30.1	26.5	1
Exercise 2: Internet	36.1	45.8	18.1	2
Exercise 3: Google Lens	27.7	24.1	48.2	3

In addition, students were asked to rank the difficulty of the different methods used to identify marine organisms, with 1 indicating the least challenging and 3 indicating the most challenging (Table 5). Based on their collective experience with all methods, students ranked the book (Exercise 1) as the least difficult method for identifying marine organisms. In contrast, they labeled the use of Google Lens (Exercise 3) as the most difficult method for accurately identifying marine organisms. These rankings provide valuable insight into the perceived ease or complexity of each identification method from the students' perspective.

Conclusion

This study highlights the effectiveness of different methods for identifying marine organisms during a workshop, as well as student preferences and experiences. Several important findings emerge from this study:

Exchangeable identification methods: Analysis of species identification revealed no significant differences among the methods tested. This suggests that different identification methods, including books, the Internet, and Google Lens, are

interchangeable without compromising the quality of species identification.

Perceived difficulty and student preferences: According to student feedback, the traditional book was the least difficult method for identifying marine organisms, while the Google Lens application presented the most difficulty. Although students found Google Lens to be more difficult, this method was preferred by them. This preference is likely due to the accessibility of Google Lens on smartphones, which makes it a versatile tool for applying knowledge gained in the classroom to real-world scenarios.

Biodiversity as Cultural Heritag: The study highlights the value of biodiversity knowledge as part of cultural heritage. Students expressed high satisfaction with their workshop experience, reflecting the positive impact of hands-on learning about marine organisms.

Even success in identification: Results from the first part of the study show that students achieved a high degree of success in identifying marine organisms using all three methods, with very similar percentages of completely correct identifications. Regardless of whether they used books, the Internet, or Google Lens, students showed consistent and impressive accuracy.

Smartphone Use Attitude: The questionnaire revealed that students had positive attitudes toward smartphone or tablet use during the workshop. They perceived smartphone use as a successful and understandable approach that was both fun and educational (Lang & Šorgo, 2023).

Satisfaction and Future Preferences: The second part of the questionnaire assessed student satisfaction and future preferences. Of note, a significant percentage of respondents indicated the least difficulty in identifying marine organisms when using books and conversely had the most challenges when using Google Lens. However, a significant percentage expressed intentions to use Google Lens in future identification activities.

The versatility of different identification methods and the educational benefits they provide, as well as the positive response of students to the use of smartphones in the learning process. It also highlights the potential of technologies such as Google Lens to extend classroom learning beyond the school environment. Overall, the findings support the notion that innovative educational approaches can enhance students' understanding of biodiversity as an integral part of cultural heritage.

References

Ahmad, Y. (2006). The scope and definitions of heritage: from tangible to intangible. *International journal of heritage studies*, *12*(3), 292-300.

Anđić, B., Cvijetićanin, S., Maričić, M., & Stešević, D. (2019). The contribution of

dichotomous keys to the quality of biological-botanical knowledge of eighth grade students. *Journal of Biological Education*, *53*(3), 310-326.

Heberlein, T. A. (2012). *Navigating environmental attitudes*. Oxford University Press.

Janžekovič, B. (2022). Insights into Biodiversity Literacy: A Proposed Model. *The American Biology Teacher*, *84*(8), 456-458.

Lang, V., & Šorgo, A. (2022). Added value of the Pl@ntnet smartphone application for the motivation and performance of lower secondary school students in species identification. *In ICERI2022 Proceedings* (pp. 4534-4540). IATED.

Lang, V., & Šorgo, A. (2023). Recognition of the Perceived Benefits of Smartphones and Tablets and Their Influence on the Quality of Learning Outcomes by Students in Lower Secondary Biology Classes. *Applied Sciences*, *13*(6), 3379.

Lowenthal, D. (2005). Natural and cultural heritage. *International Journal of Heritage Studies*, *11*(1), 81-92.

Šorgo, A., Špur, N., & Škornik, S. (2016). Public attitudes and opinions as dimensions of efficient management with extensive meadows in Natura 2000 area. *Journal of environmental management*, *183*, 637-646.

UNESCO (2023a). Cultural Heritage. Retrieved from: https://en.unesco.org/fieldoffice/santiago/cultura/patrimonio (21. 01. 2023).

UNESCO (2023b). Natural World. Heritage. Retrieved from: https://whc.unesco.org/en/natural-world-heritage/ (22. 01. 2023).

Digital Teaching Scenarios on Advanced STEM Topics for Higher Education

Emily Michailidi[1], Eilish McLoughin[2], Argyris Nipyrakis[1,3], Athanasia Kokolaki[1], Gunnar Friege[4], Priit Reiska[5], & Dimitris Stavrou[1]

[1]University of Crete, Greece, [2]Dublin City University, Ireland, [3]University of Groningen, The Netherlands, [4]University of Hannover, Germany, [5]Tallinn University, Estonia

The STEM Digitalis project's aim is to develop approaches which are effective for teaching STEM in a blended learning environment. Five STEM scenarios were identified as offering the potential to engage pre-service teachers with advanced STEM topics, to facilitate the development of their conceptualisation of STEM learning and to provide them with learning experiences which support them to develop their pedagogical content knowledge around STEM. This workshop, responding to the need for developing blended and distance learning environments for teaching advanced STEM topics in higher education, aims to familiarize participants with the educational use of various digital resources for science teaching in online and distance learning contexts and provide examples of good practices regarding teaching and learning strategies that promote meaningful use of digital technologies for teaching STEM topics in blended or distance learning environments.

Keywords: Digital tools, STEM education, higher education.

Background and Rationale

The COVID-19 pandemic forced institutions worldwide to quickly transition instruction from classrooms to configurations of distance learning. The whole enterprise was new to many academic institutions and to many staff members and students as well, with the consequence that they were not properly equipped with the necessary methodology for blended or distance education. Additional difficulties arose in the teaching of science courses that require students' hands-on interaction for them to develop not only scientific knowledge but also inquiry and practical skills for the operation of scientific equipment, data collection and processing etc. The closure of universities over a long period raised questions about the future of higher education and about the pandemic (or other crisis) preparedness of universities, in terms of whether in the future they will be in position to switch to an accessible, inclusive, equitable and engaging science education online with proper support for students and staff members (Dillon &

Avraamidou, 2020).

In this context, many efforts have been made over the last years towards developing online or blended digital resources and teaching and learning environments for tertiary science education that may be used in order to extend the capabilities of science education in cases where physical presence is not feasible. From a pedagogical perspective, online teaching has a positive impact on teaching and assessment strategies (Cook & Steinert, 2013). Teachers can incorporate effective pedagogical and instructional strategies such as games, interactive models, computer simulations and animations for learners to engage in meaningful knowledge construction. The multimodality and availability of these rich resources are advantageous as online digital resources can enrich the classrooms and improve student learning, by providing high-quality and collaborative online learning experiences synchronously and asynchronously (Eichler & Peeples, 2013).

The STEM Digitalis Project

Aims and Goals

Recognizing the aforementioned contextual needs and taking into consideration the affordances of distance or blended digital resources for science education, the STEM Digitalis Erasmus+ project (https://stemdigitalis-project.eu/) aimed at developing blended and distance learning environments for teaching advanced STEM topics to prospective primary and secondary education science teachers. The ultimate goal is to digitally equip prospective science teachers not only regarding the use of digital tools (e.g. educational robotics, data loggers, virtual and augmented reality, etc.) but also regarding the effective integration of such tools in teaching and learning processes of advanced STEM topics.

STEM Digitalis' Digital Scenarios

Five STEM scenarios were identified as offering potential to engage pre-service teachers with advanced STEM topics, to facilitate the development of their conceptualisation of STEM learning and to provide them with learning experiences which support them to develop their pedagogical content knowledge around STEM. Each scenario is presented over several units, with each unit comprising of several learning activities. These activities were designed around a five-dimensional instructional framework, depicted in Figure 1, which was developed in the STEM Digitalis project, informed by literature on integrated STEM learning (Radloff & Guzey, 2016) and blended learning (Margulieux et al., 2016). The activities within each unit were designed such that educators could choose subsets of the activities to make up a learning trajectory best suited to their programme's context. Sample trajectories were visualised using the LePlanner programme and published using the STEM Digitalis Pressbooks

platform.

Figure 1. Five-dimensional instructional framework for STEM Digitalis scenarios.



The content of each scenario was trialed in local educational contexts by each of the project members. Based on feedback from these trials, the content of the five scenarios was amended and refined. Further refinement happened following implementation of the scenarios during the STEM Digitalis Summer School (which took place in Crete in June 2022) and the STEM Digitalis Online Joint Seminars (which took place in January - September 2023).

During the ESERA 2023 workshop participants implemented three of the digital teaching scenarios developed in the context of the STEM Digitalis project, engaging with different types of educational digital tools in the context of three advanced STEM topics, namely Climate change, Water, and Ocean batteries & energy farms.

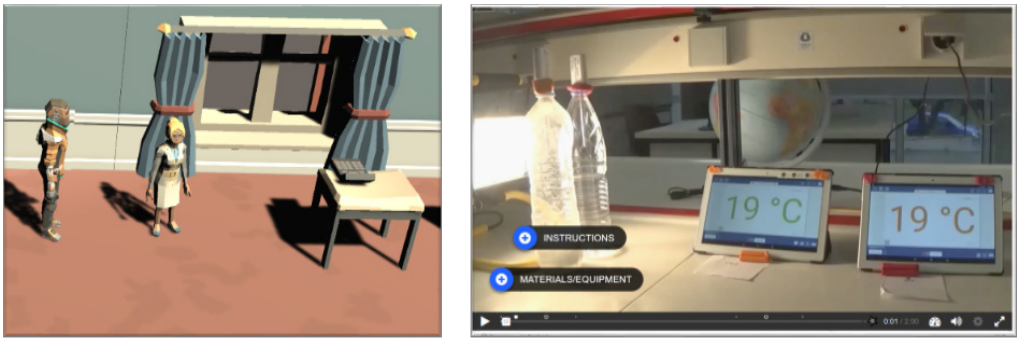
Digital scenario on Climate Change

The rapid climate change that is observed during the last century has important impacts on the natural environment, threatening the biodiversity and survival of ecosystems as well as human societies, affecting health and economies. Key leverage point to achieve the goal of citizens' climate literacy is climate education that may equip citizens with knowledge on climate change causes, consequences, and mitigation-adaptation measures. The digital teaching scenario of climate change consists of 3 units that focus on (i) drivers of climate change, (ii) carbon emissions generated by anthropogenic activities and mitigation measures taken regarding renewable sources of energy and (iii) the negotiation of a socioscientific issue regarding the lignite phase out to meet EU goals for

climate neutrality.

In order to cover these aspects of climate change diverse digital tools have been used such as: Unity 3D, mobile apps, H5P, online communication and collaboration tools. Particularly, the first unit on climate change drivers employs a Unity 3D-based serious game (shown in Figure 2a) through which participants collect and compare data concerning temperature and carbon dioxide rise through decades of 1980s until 2010s in order to form conclusions about global temperature rise and carbon dioxide average concentration rise. Participants also explore the correlation between the global temperature rise and the carbon dioxide concentration rise via an integrated interactive experiment with integrated hotspots that provide additional information, multiple choice questions and links to online dashboards for sharing conclusions, shown in Figure 2b.

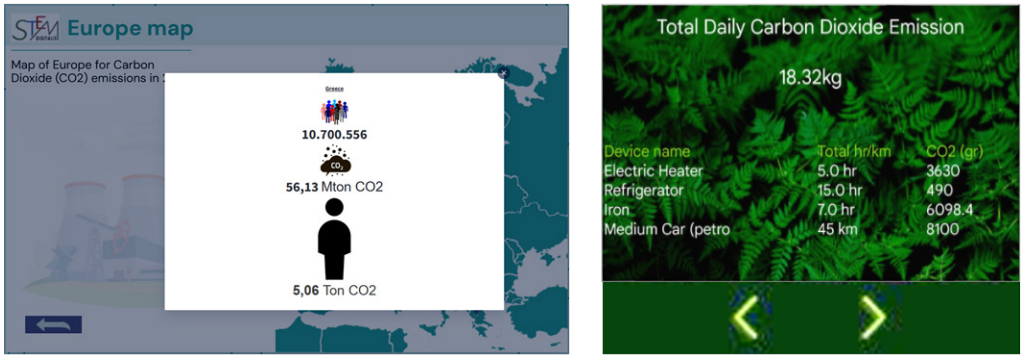
Figure 2. Snapshot of the serious game (a) and the interactive video experiment (b) environment



a b

The second unit is based on interactive maps, depicted in Figure 3a, through which participants explore data on the amounts of CO_2 emitted per anthropogenic activity per country and on a mobile application, shown in Figure 3b, through which participants calculate their personal CO_2 emissions.

Figure 3. Snapshot of the interactive map (a) and the mobile application (b)



a b

The third unit is focused on the negotiation of perspectives of multiple stakeholders on a socioscientific issue raised in the context of the need for climate neutrality. In order to do so, participants take part in a "treasure hunt" supported by an augmented reality application where they scan an image with their mobile devices and information appears on the screen. By solving riddles and using maps, participants in small groups have to trace the image of one of the stakeholders that are involved in the debate.

Digital scenario on Water

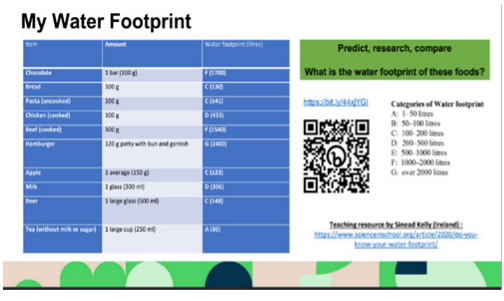
In a world faced with the demands of a growing population, increased urbanization, and climate change, the world's water resources are under threat, with Global water demand set to increase by 22% by 2030. This highlights the urgent need for research, and the finding of technological solutions to combat our growing water problems, national and internationally. Sustainable water means a nation that can be water self-sufficient: ensuring there is enough water to meet multiple needs, from agriculture to municipal and industrial. This scenario focuses on developing learners' understanding of how individuals and societies need to work together to provide safe, secure drinking water for all. These resources and activities help us to think about how we use water and how we might use a little less.

The digital teaching scenario of water consists of 3 units that focus on developing understanding of (i) water conservation and the water cycle, (ii) an individual's water footprint and the role of citizen science in water education and (iii) water monitoring and data analysis. This scenario promotes the use of digital communication and collaboration tools, CODAP web-based software for data analysis, CMAP Concept mapping software, mobile apps, online forms and databases.

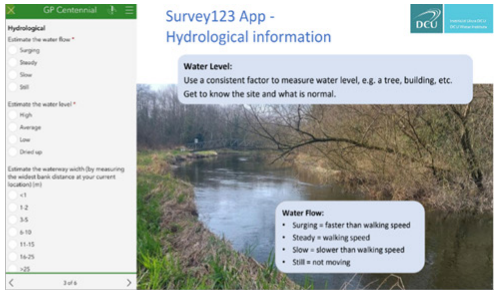
The first unit facilitates learners to prepare causal explanations and visual

representations of the water cycle. As shown in Figure 4a, learners calculate and reflect on their own water footprint based on daily food consumption.

Figure 4. Snapshots of water footprint activity (a) and water quality activity (b)



a

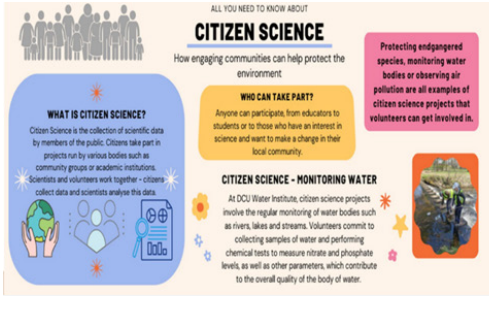


b

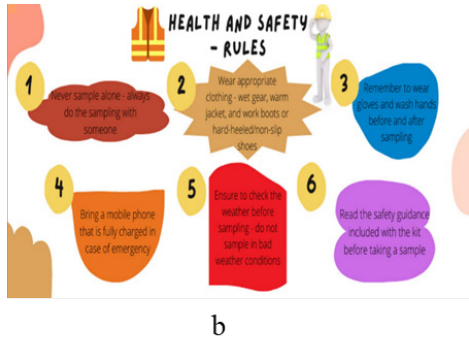
The second unit focuses on the notion of data literacy and its importance in terms of social justice in the 21st century as well as on the development of skills in graphing and interpreting water data sets and observations. Particularly learners are presented with a large dataset, recorded via a multi-sensor sonde in Dublin port and asked to interrogate the dataset using CODAP.

The third unit introduces learners to the opportunities and benefits of being involved in co-created citizen science projects around water, as shown in Figure 5. Specifically, learners work in groups to collect water samples, to test them and to upload their findings to the Freshwater Watch project database.

Figure 5. Snapshots of the citizen science activities



a



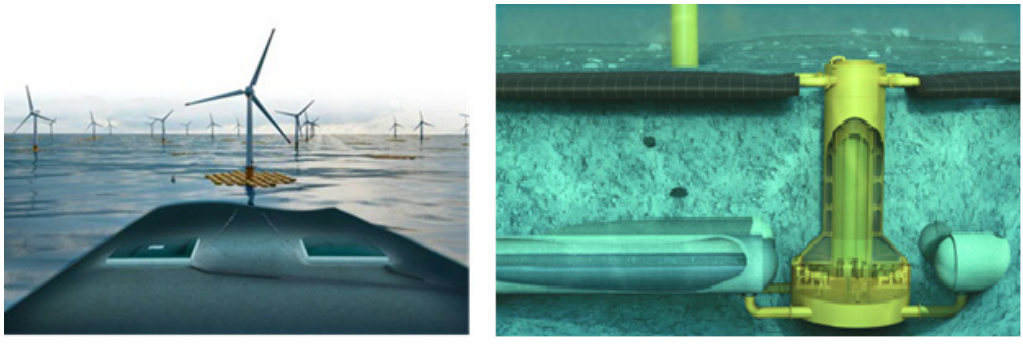
b

Digital scenario on Ocean Batteries & Ocean Energy Farms

The growing need for sustainable and renewable energy sources appears as an eco-friendly solution to the energy crisis which is a contemporary global problem. However, fluctuations appear in the energy supply from renewable energy sources due to the instability and unpredictability of local weather

conditions (e.g. wind, sunlight). Therefore, it appears an additional need for sustainable and eco-friendly energy storage techniques that would support the operation of renewable energy farms. In parallel, research on off-shore energy farms consisting of wave energy converters and wind turbines aims for higher efficiency levels in energy harvesting. In light of the above, Ocean Grazer (www.oceangrazer.com) is a hybrid renewable energy harvester, which means that it integrates both energy production and energy storage modules (DiModugno, 2021). Indicative illustrations of such a module are presented in Figure 6. Specifically, the installation combines wave energy converter technology (Asiikkis et al., 2023) with wind turbines and on-site energy storage devices called 'ocean batteries' in order to harvest energy off-shore (Vissers, 2019). Ocean batteries are devices that store energy in the form of potential energy by making use of differences in hydrostatic pressure between an internal rigid reservoir kept in low (atmospheric) pressure and a flexible bladder kept in high pressure at the bottom of the installation inside the ocean.

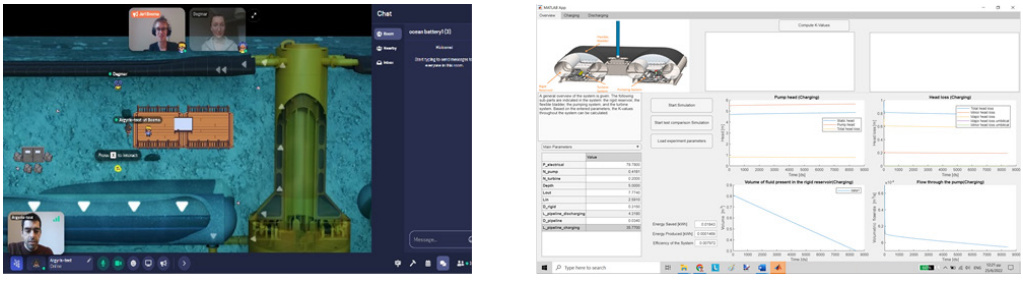
Figure 6. Illustrations from the Ocean Grazer μodule (oceangrazer.com)



This teaching module, based on the aforementioned phenomena and applications, engages students with engineering practices on exploring, testing, revising, and reflecting on a simulation model on ocean batteries and ocean energy farms, while they collaborate with peers through an online gamification platform. In specific, the module comprises three submodules. In the first submodule, students log into a customised gamification environment (gather.town) as shown in Figure 7a, formulate small groups online, and go through a historical overview of renewable energy sources. In order to proceed among stages in the platform, the users interact with virtual assistants through Augmented Reality mobile apps (Metaverse app), who also assess their knowledge in each stage in order to unlock the passwords needed to proceed to the next stage. In the second submodule, users engage in engineering design practices by carrying out several variable testing procedures through a Matlab-based simulation programme (Ocean Grazer), as shown in Figure 7b, in order to optimise the

design of i) ocean batteries (charging and discharging features), ii) wave energy converters, iii) wind turbines. In the third submodule, students reflect on their design decisions on the ocean batteries and energy farm module, as well as on interdisciplinary, epistemological, and socioscientific aspects related to the activities they previously experienced.

Figure 7. Snapshot of the gamification platform (a) and of the ocean battery simulation (b)



a b

Key learnings from digital scenarios' implementation

Based on the implementation of the digital scenarios, engaging participants in student-centred and inquiry-based approaches emerges as crucial to integrating technology for STEM teaching and learning. The design and implementation of the scenarios provided some exemplar ways in which technology can be used as an indispensable part of the teaching scenario e.g., in order to collect and analyse data and draw inferences that can lead to achieving the teaching goals.

Moreover, technology can be used as a complementary asset of the educator by aiding in an individualised and flexible manner, e.g., by providing content- or process-related assistance and feedback to participants to accomplish the tasks. However, assistance and guidance from the teacher were deemed important in the implementation of the scenarios; hence we can conclude that a combination of assistance from both educators and technology can be the most productive method.

Additionally, the implementation of contemporary and real-world topics can provide a fruitful context that can stimulate participants' interest and engagement with the content, phenomena, and applications. In this light, taking into consideration sustainability goals such as aiming at affordable and clean energy and climate action can activate meaningful integration of content topics that relate to participants' everyday life. Also, engaging participants with socioscientific issues that are relevant to the topic can activate an additional metalevel of participants' understanding and engagement with the topic.

However, discussions about the impact of technologies are recommended to be assisted by educators and experts in order to discuss more deep insights about the topic.

Finally, collaboration with peers, educators and experts is an important factor that can enhance learning and engagement with the topic, as well as it can assist in overcoming difficulties in using technology. However, adequate interaction across different groups of participants is needed to take into consideration multiple views. In this direction, technological tools and platforms can provide an environment that can assist collaboration in distance and blended learning environments.

References

Asiikkis, A., Grigoriadis, D., & Vakis, A. (2023). Maximizing the Surge Amplitude of a Floater Through an Adaptable Mooring Tightening Technique. *Proceedings of the European Wave and Tidal Energy Conference, 15*. https://doi.org/10.36688/ewtec-2023-274

Cook, D. A., & Steinert, Y. (2013). Online learning for faculty development: A review of the literature. *Medical Teacher, 35*(11), 930–937. https://doi.org/10.3109/0142159X.2013.827328

Di Modugno, A. (2021). *A Techno-Financial Model of Ocean Grazer's Hybrid Integration Proposing the Ocean Battery as Deployable On-site Storage Profitability Optimizer for Offshore Renewable Production Plants.* [Master Design project, University of Groningen].

Dillon, J., & Avraamidou, L. (2020). Towards a viable response to COVID-19 from the science education community. *Journal for Activist Science & Technology Education, 11*(2), 1–6. https://doi.org/10.33137/jaste.v11i2.34531

Eichler, J. F., & Peeples, J. (2013). Online homework put to the test: A report on the impact of two online learning systems on student performance in general chemistry. *Journal of Chemical Education, 90*(9), 1137–1143. https://doi.org/10.1021/ed3006264

Margulieux, L.E., McCracken, W.M. & Catrambone, R., (2016). A taxonomy to define courses that mix face-to-face and online learning. *Educational Research Review, 19*, 104-118. https://doi.org/10.1016/j.edurev.2016.07.001

Radloff, J. & Guzey, S. (2016). Investigating pre-service STEM teacher conceptions of STEM education. *Journal of Science Education and Technology, 25*, 759-774. https://doi.org/10.1007/s10956-016-9633-5

Vissers, V. (2019). *Energy storage in a world powered by renewables.* [Bachelor Integration project, University of Groningen].

Using Augmented Reality and Simulations to Support Learning of Fundamental Electrical Concepts

Florian Frank, Christoph Stolzenberger, Markus Elsholz and Thomas Trefzger

University of Wuerzburg, Germany

Secondary school students struggle to fully understand fundamental electrical concepts such as the relation and difference between current and voltage. According to multimedia learning theories, the use of augmented reality applications or simulations can support students in this process by helping to manage their cognitive load while learning. A study was conducted with 177 secondary school students, investigating the effects of the use of a simulation and an Augmented Reality (AR)-application for supporting the students while learning about electricity. The use of the simulation and the AR-application was compared with the use of traditional informational graphics for presenting analogy-based models for electricity. Over the course of a student laboratory consisting of four units pertaining to fundamental electrical concepts (electrical current and voltage, electrical resistance, parallel and series circuits), the cognitive load the students perceived was measured. The questionnaire used was examined via a confirmatory factor analysis, after which the data was processed according to an Item-Response-Theory-approach. Subsequent analyses showed that students reported a significantly lower extraneous cognitive load when working with the AR-application as compared to working with the simulation, but no significant differences when compared to working with informational graphics.

Keywords: Physics, multimedia-supported learning, cognitive load.

Introduction

Recent studies of student's misconceptions about electricity have shown that after the introductory lessons on simple electrical circuits, almost two thirds of students have difficulties correctly separating between the concepts of electrical current and voltage (Ivanjek et al., 2021), with more than half the students misunderstanding voltage to be a property of electrical current (Burde, 2018, p. 244). These misconceptions are still prevalent after finishing early secondary school (Müller et al., 2022) and can even be found among first semester university students (Rahmawati et al., 2023). These empirical findings suggest that the way electricity is taught in schools may not be sufficient to enable the students to fully understand electricity and its related core concepts. To enhance

the students' understanding, teachers frequently use analogy models based on systems the students are more familiar with like air pressure (Burde & Wilhelm, 2020). While the use of these models can support learning, the way the models are presented is crucial, since weaknesses in model presentation can lead to further difficulties and inhibit correct understanding (Mogstad & Bungum, 2020)

Based on this analysis, a research project was conducted which aimed to study and thereby gain a more accurate understanding of the difficulties of presenting didactic models for electricity. Traditionally, these models have been taught to the students via the use of informational graphics, which challenged students with the cognitively demanding task of relating the static images to the real-life circuits they are trying to understand. To ease this cognitive burden, we developed digital support material in the form of an Augmented Reality (AR)-application and a simulation, which are designed to present adaptive digital representations of didactic models for electricity fitting precisely to the real-life circuits the students are studying. Findings from Altmeyer et al. (2020) and Thees et al. (2020), relating to the use of AR-applications in learning about electricity, point towards the capacity of AR-applications to improve learning. Altmeyer et al. (2020) report a higher learning gain for AR with equal cognitive load as compared to a non-AR-intervention, whereas Thees et al. (2020) showed an equal learning gain with lower extraneous cognitive load for the AR-intervention. The effects of using the digital support material developed by us on the learning process were evaluated in an experimental study taking place at the student lab. The full research project encompasses the analysis of the student's conceptual knowledge growth, of the cognitive burden placed on the students, and of the time spent on the learning tasks. In this article, we focus on the cognitive load experienced while learning and how it is influenced by the use of multimedia technology (simulations and AR).

Theory and Predictions

Cognitive Load Theory

A central tenet of the *Cognitive Load Theory (CLT)* is that every student's cognitive capacity is finite (Sweller et al. 2019). This cognitive capacity is utilised in learning, where every learning task imposes a certain cognitive load. If that cognitive load exceeds the cognitive capacity, the learning process is hindered. One can distinguish between three categories of cognitive load: intrinsic cognitive load (ICL), extraneous cognitive load (ECL) and germane cognitive load (GCL). Intrinsic load results from processes contributing to an understanding of the learning content and is therefore independent of the learning situation or mode of content presentation. Processes aimed at understanding the presentation of the content result in extraneous load. Germane load is generated

by processing the learning content, storing it in long-term memory and linking it to previously acquired information. Following this, an ideal learning task should primarily impose intrinsic and germane cognitive load, since these are caused by understanding and processing the learning content, and should avoid imposing extraneous load, since it is a product of trying to understand not the learning content, but the presentation thereof.

Cognitive Theory Of Multimedia Learning

The *Cognitive Theory of Multimedia Learning (CTML)*, which adapts these principles of the CLT for complex learning scenarios involving multiple sources of information, contains design guidelines that aim to reduce the overall cognitive load imposed by the learning situation (Mayer, 2014). Following the principles of segmentation and coherence, the learning content should be separated into the small units, containing only the necessary information, thereby reducing the intrinsic load imposed to the minimum that is necessary (Mayer & Moreno, 2010). Regarding the extraneous load, according to the principles of spatial and temporal contiguity, the ECL can be reduced by making all resources relevant to the learning process available at the same time and in the same place, thereby reducing the need for taxing use of short-term memory storage (Mayer & Fiorella, 2014).

Key Considerations And Hypotheses

We propose that the central resources relevant to a learning process regarding electricity are the real-life experiment (e.g. a simple circuit with a light bulb) and the didactic model based on an analogy to understand the mechanisms imperceptible in the experiment (e.g. a model based on a height analogy explaining how the battery causes an electrical current, also see Mogstad & Bungum, 2020). Since the resources are partly real/observable (experiment) and partly theoretical/imaginary (the didactic model), the presentation of both at the same time and in the same place is not possible while using informational graphics to convey the didactic model. However, a simulation might be used to superimpose the model atop a digital replica of the circuit, thereby combining both resources, albeit in a solely digital environment. The use of a see-through-AR-application allows the presentation of the model atop the real-life circuit, theoretically strengthening the connection between the experiment and the model even further.

We designed a student lab program consisting of four units of increasing complexity and internal difficulty to study the effects of the use of a simulation and an AR-application on the perceived cognitive load compared to using informational graphics to present the didactic models. Our hypotheses were the following:

H1: The intrinsic cognitive load does not differ between the three modes of

model presentation.

H2: The intrinsic cognitive load differs between the four units.

H3: The extraneous cognitive load differs between the three modes of model presentation.

H4: The germane cognitive load differs between the three modes of model presentation.

Methods

We used an experimental design with multiple tests to compare a traditional approach (intervention group IG) to present didactics models with a simulation-based approach (interventions group SIM) and an AR-based approach (intervention group AR) at four different complexity levels.

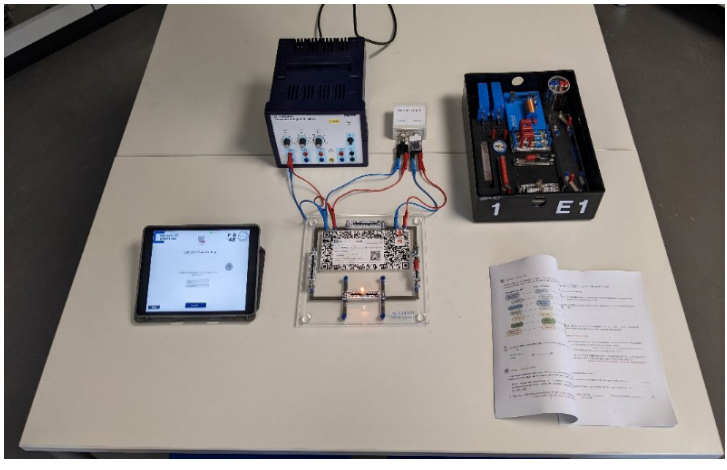
Participants

The investigations were carried out in 2022 and 2023. The participants attended local eight-grade secondary school classes. Before the students took part in the intervention, they had already completed their lessons on electricity and its basic concepts (electrical current, voltage, and resistance). In total, N = 177 8th-grade students participated and were randomly assigned to an intervention group (N = 67 in group IG, N = 64 in group SIM, and N = 46 in group AR).

Intervention Design

The intervention took place in an out-of-school student laboratory at the University of Würzburg. In the student lab, the participants completed four units in a fixed order, starting with a unit on "Electrical Current and Voltage". This unit was designed to primarily contain previously taught information to give the students an opportunity to familiarize with the student lab and the digital support materials they would be working with. The next unit, "Electrical Resistance", involved new content and unfamiliar tasks, e.g. using the models presented in the digital support material to explain simple observations. Units 3, "Parallel Circuits", and 4, "Serial Circuits", required the participants to use the models presented in the digital support material to form hypotheses about Kirchhoff's circuit laws, involving complex reasoning and justification tasks. During units 1 to 4, the students used two didactic models of electricity (based on a height analogy of a marble track and based on an air pressure analogy, respectively) to learn about electricity.

Figure 1. Material used in the student lab (from left to right: tablet computer hosting the digital support materials, experimentation set for simple circuits, printed-out workbook)



The students worked in small groups, supervised by university student teachers. The students were supplied with a printed-out workbook, containing the instructions and tasks described above, as well as an experimentation set for simple circuits and a tablet computer (see Figure 1). The structure and contents of the workbook are the same across all intervention groups. At multiple points in the workbook, the students are instructed to use the tablet to access presentations of the didactic models. When developing the intervention, care was taken to ensure that all groups were asked to interact with the digital support material at the same time to equalize the motivational effects of technology use across all groups as much as possible. After each unit, the students were asked to fill out a short test, followed by a 10–15-minute rest before starting with the next unit.

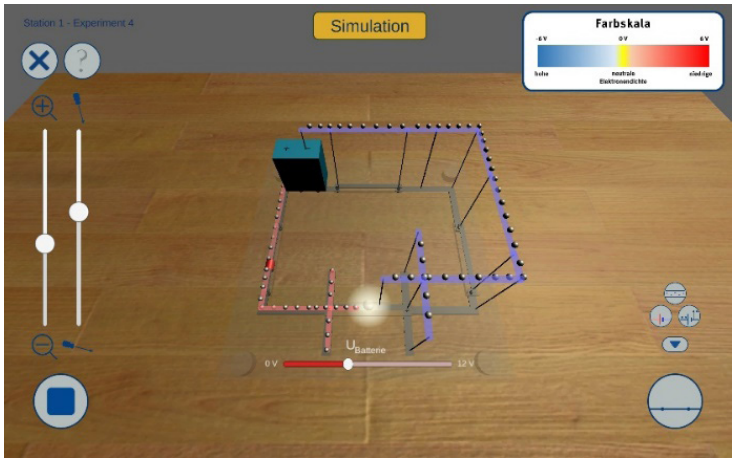
Differences Between Intervention Groups

The traditional visualization of the models via infographics (intervention group IG) was provided to the students via the online platform *tet.folio* (Haase, Kirstein & Nordmeier, 2016). To that end, a tet.book was created that contained informational graphics taken from ordinary schoolbooks used in Germany. The graphics were in part supported by short annotations. The goal of the tet.book called "Visualisations" was to present the didactic models in a concise and familiar way.

To present the didactic models for intervention group SIM, a simulation was developed, that contained a three-dimensional replica of the real experimentation set used in the student lab. The students recreate the real-life circuit they had built via the simulation by dragging the respective components to their spot. After the

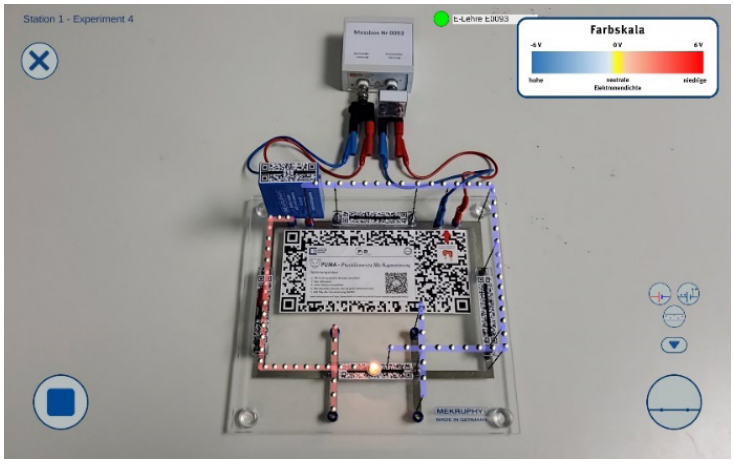
construction of the virtual circuit is completed, the students use the simulation to visualize the didactic models, which were fitted to the digital circuit and adapted to changes in the digital circuit (e.g. when closing or opening a switch) in real time (see Figure 2).

Figure 2. Simulation-based approach to presenting the didactic models for electricity (shown are an air-pressure- and a height-based model in the context of a simple circuit with one light bulb)



For intervention group AR, a similar application has been developed, using augmented-reality-technology to overlay the didactic models atop the real-life circuit (Stolzenberger et al., 2022). It works in the same way as the simulation described above: after building a circuit, you use the camera of the tablet computer to scan the experiment. On successful scan, the application then visualizes the didactic models by superimposing digital representations of them onto the real-time camera feed of the tablet computer (see Figure 3). Using a small Bluetooth-measurement box, a connection between the circuit and the AR-application is established, which again allows for changes made to the circuit to impact the digital representation of the didactic models.

Figure 3. AR-based approach to presenting the didactic models for electricity (shown are an air-pressure- and a height-based model in the context of a simple circuit with one light bulb)



Dependent Variables And Test Instruments

Cognitive Load

The cognitive load was measured using the Cognitive Load Scale (Klepsch et al., 2017). The scale contains 8 items, 2 pertaining to measuring the intrinsic load, 3 measuring the extraneous load, and 3 measuring the germane load. The authors themselves however suggest only using 2 of the 3 germane load-items, since one item is "only useful if GCL is varied on purpose in a given learning material (e.g., through prompts)" (Klepsch et al., 2017, p. 10). Every item contains a statement about the learning task, the students are asked to note their agreement or disagreement on a 7-point Likert scale.

The scale in its entirety (all 8 items) was administered during the short post-test after every unit, in total 4 times.

Conceptual Knowledge and Time on Task

The conceptual knowledge development was measured using the 2T-SEC Test developed by Ivanjek et al. (2021). It was administered in the pre-test, together with the measurement of the academic achievement of the students (self-reported grades in mathematics and science), and in the four post-unit-tests, together with the cognitive load scale. For the post-tests, the original test was divided evenly into four parts pertaining to the contents of the four units.

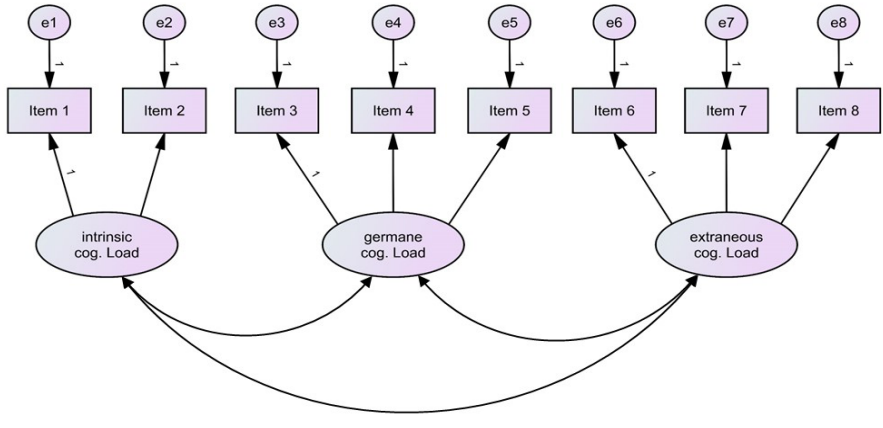
The Time-on-Task was logged by the teaching assistants, totaling to four separate Time-on-Task-measurements per every student.

Results

Preparing The Cognitive Load Data

We started by examining the structure of the Cognitive Load Scale. Although the material used by the intervention groups did not differ in the number of prompts or similar didactic scaffolding, we purposefully used the full 8-item scale with the intent of using a confirmatory factor analysis to support the authors claim of leaving out one item pertaining to the germane load.

Figure 4. Structure of the Cognitive Load Scale (full 8-item-variation, the 7-item-variation removes item 5)



SPSS AMOS was used to model the structure (see Figure 4) and compare the two versions of the instrument. The structure was confirmed for each post-unit-test separately, so four times in total. The model fit was evaluated using the probability of obtaining the χ^2-Value (higher χ^2-P indicates better fit), the comparative fit index (higher CFI indicates better fit), the root-mean-square error of approximation (lower RMSEA indicates better fit) and the standardized root mean squared residual (lower SRMR indicates better fit). Following Hu & Bentler (1999), an acceptable fit would result in the CFI exceeding .95, with the RMSEA and the SRMR staying below .06 and .08, respectively. Comparing the model fit indices obtained by our analysis of the cognitive load scale, we concluded the model containing 7 items to be the better fitting model (see Table 1 for details). Although the 7-item-variation fits our data better, it must be noted, it did not result in only passable fit indices. While the SRMR showed acceptable values, only half of the past-unit-models achieved a CFI of above .95 and none

reached a RMSEA-value of below .06. Klepsch et al. (2017) analyzed the 7-item-variation as well and reported a χ^2-P-value of below .001, a CFI-value of .97 and a RMSEA-value of .021. Further analysis will be conducted to investigate this discrepancy between their reported fit indices and the fit indices found for our data.

Table 1. Model Fit of the two Cognitive Load Scale Variations (the SRMR-value of the 7-item-variation for Unit 1 could not be computed by SPSS AMOS).

	8-item-variation				7-item-variation			
	χ^2-P	CFI	RMSEA	SRMR	χ^2-P	CFI	RMSEA	SRMR
Unit 1	**< .001**	**.829**	**.121**	**.0925**	**< .001**	**.908**	**.101**	*
Unit 2	**< .001**	**.903**	**.109**	.0757	< .001	.921	.113	.0675
Unit 3	< .001	.920	.110	.0802	.005	.969	.077	.0492
Unit 4	< .001	.913	.124	.0846	< .001	.955	.100	.0686

To prepare the Cognitive Load Data for further analysis, the R-Package TAM (Robitzsch et al., 2022) was used to process the raw data and scale it according to an Item-Response-Theory-Approach (Embretson & Reise, 2009, p. 40-61). Since the Cognitive Load Scale splits into three separate subscales pertaining to each of the types of cognitive load, we used three separate IRT-models to process the data. Each of these models contained the repeated measures Likert-data for one of the cognitive load types, resulting in a four-dimensional (for the four post-unit tests) Generalized Partial Credit (to process the Likert-data) model (Embretson & Reise, 2009, p. 110-115). Preparing the data this way results in factor scores representing the perceived cognitive load, where the higher the factor score, the higher the perceived cognitive load.

Analysis of Group-Wise Intrinsic Load Data

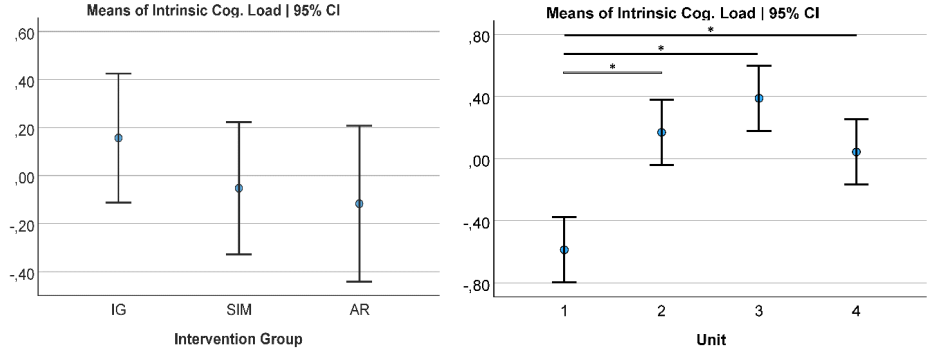
To interpret the differences between the intervention groups, we calculated the mean intrinsic cognitive load experienced across the four units. Afterwards, a One-Way-ANOVA was conducted to investigate the effect of the model presentation on this mean intrinsic cognitive load. We found no statistically significant main effect of intervention group (and therefore model presentation) on the mean intrinsic cognitive load, $F(2,174)=.980$, $p=.377$, partial $\eta^2=.011$ (see Figure 5).

Analysis Of Unit-Wise Intrinsic Load Data

We conducted a One-Way-ANOVA to investigate the effect of the content learned during the different units on the experienced intrinsic cognitive load. We found a statistically significant main effect of the unit (and therefore the learning content) on the intrinsic cognitive load, $F(3,700)=15.409$, $p<.001$,

partial η^2=.062. Post-Hoc Tests indicated that the mean intrinsic cognitive load in unit 1 (M=-.586, SD=.107) was significantly different (p<.05) from the mean intrinsic cognitive load in unit 2 (M=.169, SD=.107), unit 3 (M=.388, SD=.107) and unit 4 (M=.043, SD=.107). No other statistically significant differences in intrinsic cognitive load were found (see Figure 6).

Figure 5 & 6. Means of the Intrinsic Cognitive Load by model presentation (on the left) and by unit (on the right)



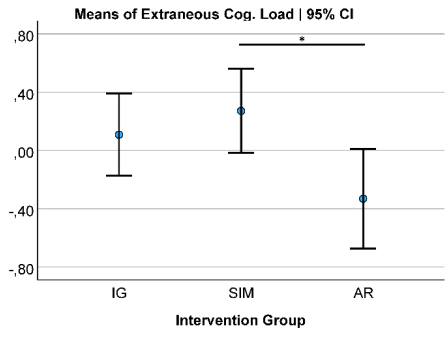
Analysis Of Group-Wise Extraneous Load Data

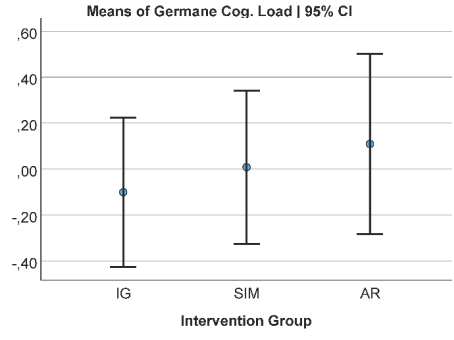
To interpret the differences between the intervention groups, we calculated the mean extraneous cognitive load experienced across the four units. Afterwards, a One-Way-ANOVA was conducted to investigate the effect of the model presentation on this mean extraneous cognitive load. We found a statistically significant main effect of intervention group (and therefore model presentation) on the mean extraneous cognitive load, F(2,174)=3.663, p=.028, partial η^2=.040. Post-Hoc Tests indicated that the mean extraneous cognitive load in intervention group AR (M=-.332, SD=.173) was significantly different (p<.05) from the mean extraneous cognitive load in intervention group SIM (M=.271, SD=.146). The mean extraneous cognitive load in intervention group IG (M=.108, SD=.143) did not significantly differ from the mean extraneous cognitive load in intervention groups SIM (p=.811) and AR (p=.147) (see Figure 7).

Analysis Of Group-Wise Germane Load Data

To interpret the differences between the intervention groups, we again calculated the mean germane cognitive load experienced across the four units. Afterwards, a One-Way-ANOVA was conducted to investigate the effect of the model presentation on this mean germane cognitive load. We found no statistically significant main effect of intervention group (and therefore model presentation) on the mean germane cognitive load, F(2,174)=.336, p=.715, partial η^2=.004 (see Figure 8).

Figure 7 & 8. Means of the Extraneous Cognitive Load by model presentation (on the left) and the Germane Cognitive Load by model presentation (on the right)





Discussion and Conclusion

Regarding The Intrinsic Cognitive Load

The intrinsic cognitive load did not differ between the modes of model presentation, hypothesis H1 is therefore accepted. This finding supports our claim that the content the students learned during the intervention did not differ between the different modes of model presentation. Great care was taken when designing the intervention to ensure this being the case, since the goal of our research was to investigate different modes of presenting and learning the same content.

The unit-wise analysis revealed a difference in intrinsic cognitive load between unit 1 and units 2, 3 & 4, hypothesis H2 is therefore accepted. As previously stated, Unit 1 focused on familiarizing the participants with the student lab and the material, therefore primarily containing already taught content. Units 2 to 4 differed from this and confronted the students with new content and tasks. This finding has been described as the Pre-Training-Principle, which states that when learning, the intrinsic cognitive load can be reduced by using previously stored knowledge to help chunk the incoming information (Mayer & Moreno, 2010, p. 145-146.

Regarding The Extraneous Cognitive Load

It was hypothesized that the use of multimedia for presenting the models would result in a difference in extraneous cognitive load. Although a significant difference between the modes of model presentation was found, leading us to accept hypothesis H3, it was expected to be a significant reduction when using the simulation and the AR-application. Instead, the AR-application induced significantly lower extraneous cognitive load as compared to the simulation, while the use of traditional informational graphics did not significantly differ

from these two modes of presentation. We propose the following interpretation of this finding: The central learning resources are, as stated previously, the real-life experiment and the intangible didactic model. The AR-application manages to combine both by superimposing a representation of the didactic model atop the real-life experiment. The simulation tries to achieve a combination of both resources by substituting the real-life experiment with a digital one, thereby introducing a new source of content presentation not present in the AR-application. This likely lead to the students having to process more simultaneous content presentation, thereby raising their perceived extraneous cognitive load. Although we found a difference in mean extraneous cognitive load between intervention groups IG and AR, it was not statistically significant.

Regarding The Germane Cognitive Load

The analysis revealed there to be no significant difference in mean germane cognitive load between the three modes of model presentation, hypotheses H4 is therefore rejected. Initially, the alternative hypothesis was chosen, because we presumed that learning about electricity overburdens the students cognitively. If the extraneous cognitive load while learning could be reduced, capacity would be freed up to be used for processing and memorizing the learning content, which would result in a higher germane cognitive load. Since the different modes of model presentation were presumed to yield different extraneous cognitive loads, one would then also expect a difference in germane cognitive load. Because we found no difference in germane cognitive load, we conclude that at least over the course of the student lab, we did not cognitively overburden the participants.

Conclusion

In an experimental study taking place at an out-of-school student laboratory, research was conducted to investigate the effect the use of multimedia technology for presenting analogy-based models for electricity (in the form of a simulation and an AR-application) has on the cognitive burden induced by the learning process. It was found that the use of an AR-application can significantly reduce the extraneous cognitive load caused by the processing of the content presentation as compared to the use of a simulation. Using informational graphics to present the models resulted in a non-significant difference in extraneous cognitive load compared to the use of the simulation and the AR-application. Our cumulative findings affirm known principles of the *Cognitive Theory of Multimedia Learning*, more precisely the Principles of Temporal and Spatial Contiguity (pertaining to the extraneous cognitive load, shown by the reduction of the load when using the AR-application, Mayer & Fiorella, 2014) and the Pre-Training-Principle (pertaining to the intrinsic cognitive load, Mayer & Moreno, 2010). Meta-Reviews by Radu (2014) and Buchner et al. (2022) report

the use of AR-application in learning to have a positive effect on knowledge gain, group collaboration, and perceived cognitive load when compared to non-AR-learning. Radu (2014) also reports that AR-applications can overburden learners if implemented non-optimally, which is akin to our finding for the use of a simulation for model presentation.

The results presented here are preliminary, the study is currently ongoing and will in total include approximately N = 360 participants. Future analyses will also include the data on the conceptual knowledge development and Time-on-Task, as well as constructs presumed to be moderating variables (i.e. the technological affinity or the visualization capabilities of the participants).

The materials used in the presented research (including the AR-application and the simulation) are made available for free at https://go.uniwue.de/puma-s. Additionally, the multimedia applications have been published in the Apple App-Store and the Google Play Store under the name "PUMA : Spannungslabor".

Acknowledgement

This project is part of the "Qualitätsoffensive Lehrerbildung", a joint initiative of the Federal Government and the Länder which aims to improve the quality of teacher training. The programme is funded by the Federal Ministry of Education and Research. The authors are responsible for the content of this publication.

References

Altmeyer, K., Kapp, S., Thees, M., Malone, S., Kuhn, J., & Brünken, R. (2020). The use of augmented reality to foster conceptual knowledge acquisition in STEM laboratory courses—Theoretical background and empirical results. *British Journal of Educational Technology*, *51*(3), 611–628. https://doi.org/10.1111/bjet.12900

Buchner, J., Buntins, K., & Kerres, M. (2022). The impact of augmented reality on cognitive load and performance: A systematic review. *Journal of Computer Assisted Learning*, *38*(1), 285–303. https://doi.org/10.1111/jcal.12617

Burde, J.-P. (2018). *Konzeption und Evaluation eines Unterrichtskonzepts zu einfachen Stromkreisen auf Basis des Elektronengasmodells*. Logos Verlag Berlin.

Burde, J.-P., & Wilhelm, T. (2020). Teaching electric circuits with a focus on potential differences. *Physical Review Physics Education Research*, *16*(2), 020153. https://doi.org/10.1103/PhysRevPhysEducRes.16.020153

Embretson, S. E., & Reise, S. P. (2009). *Item response theory for psychologists* (Reprinted 2009 by Psychology Press). Psychology Press.

Haase, S., Kirstein, J., & Nordmeier, V. (2016). The Technology Enhanced Textbook: An HTML5-based Online System for Authors, Teachers and Learners. In L.-J. Thoms & R. Girwidz (Eds.), *Selected Papers from the 20th International*

Conference on Multimedia in Physics Teaching and Learning (pp. 85-92). Mulhouse: European Physics Society.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Ivanjek, L., Morris, L., Schubatzky, T., Hopf, M., Burde, J.-P., Haagen-Schützenhöfer, C., Dopatka, L., Spatz, V., & Wilhelm, T. (2021). Development of a two-tier instrument on simple electric circuits. *Physical Review Physics Education Research*, *17*(2), 020123. https://doi.org/10.1103/PhysRevPhysEducRes.17.020123

Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and Validation of Two Instruments Measuring Intrinsic, Extraneous, and Germane Cognitive Load. *Frontiers in Psychology*, *8*, 1997. https://doi.org/10.3389/fpsyg.2017.01997

Mayer, R. E. (2014). Cognitive Theory of Multimedia Learning. In R. E. Mayer (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (2. Aufl., S. 43–71). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.005

Mayer, R. E., & Fiorella, L. (2014). Principles for Reducing Extraneous Processing in Multimedia Learning: Coherence, Signaling, Redundancy, Spatial Contiguity, and Temporal Contiguity Principles. In R. E. Mayer (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (2. Aufl., S. 279–315). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.015

Mayer, R. E., & Moreno, R. (2010). Techniques That Reduce Extraneous Cognitive Load and Manage Intrinsic Cognitive Load during Multimedia Learning. In J. L. Plass, R. Moreno, & R. Brünken (Hrsg.), *Cognitive Load Theory* (1. Aufl., S. 131–152). Cambridge University Press. https://doi.org/10.1017/CBO9780511844744.009

Mogstad, E., & Bungum, B. (2020). Ski lifts, bowling balls, pipe system or waterfall? Lower secondary students' understanding of analogies for electric circuits. *Nordic Studies in Science Education*, *16*(1), 37–51. https://doi.org/10.5617/nordina.6882

Radu, I. (2014). Augmented reality in education: A meta-review and cross-media analysis. *Personal and Ubiquitous Computing*, *18*(6), 1533–1543. https://doi.org/10.1007/s00779-013-0747-y

Rahmawati, R., Widiasih, W., Marisda, D. H., & Riskawati, R. (2023). Using Four-Tier Test to Identify Prospective Elementary Teacher Students' Misconception on Electricity Topic. *Jurnal Penelitian Pendidikan IPA*, *9*(10), 7793–7802. https://doi.org/10.29303/jppipa.v9i10.3272

Robitzsch, A., Kiefer, T., & Wu, M. (2022). *TAM: Test Analysis Modules*. https://CRAN.R-project.org/package=TAM

Stolzenberger, C., Frank, F., & Trefzger, T. (2022). Experiments for students with built-in theory: 'PUMA: Spannungslabor' – an augmented reality app for studying electricity. *Physics Education*, *57*(4), 045024. https://doi.org/10.1088/1361-6552/ac60ae

Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Thees, M., Kapp, S., Strzys, M. P., Beil, F., Lukowicz, P., & Kuhn, J. (2020). Effects of augmented reality on learning and cognitive load in university physics laboratory courses. *Computers in Human Behavior*, *108*, 106316. https://doi.org/10.1016/j.chb.2020.106316

Wilhelm, T., Müller, S., & Burde, J.-P. (2015). Vergleich von Schülervorstellungen zur Elektrizitätslehre in Hessen und Weißrussland. *PhyDid B - Didaktik der Physik - Beiträge zur DPG-Frühjahrstagung*, *0*(0). https://ojs.dpg-physik.de/index.php/phydid-b/article/view/629

Characterising The Educational Potentials of Citizen Science Mobile Applications Related to Biodiversity

Giuliana Morbidoni-Davicino[1], María Carla Lábaque[2] and Leticia Garcia-Romano[1,2]

[1]Science and Technology Teaching Department, Faculty of Exact, Physical and Natural Sciences, National University of Córdoba, Argentina

[2]National Scientific and Technical Research Council (CONICET), Faculty of Exact, Physical and Natural Sciences, National University of Córdoba, Argentina

We characterised twenty free-download citizen science mobile applications (apps) aimed at teaching and learning about biodiversity conservation. The analysis includes characteristics related to multimedia, content, citizen engagement, and collaborative learning. The results show that there are diverse areas of knowledge on biodiversity where citizen science can be developed and that multimedia resources favour the successful identification of organisms. In addition, several apps allow citizens not only to collect data but also to participate in scientific projects, thus favouring collaborative learning and recognition of their immediate environment. It is recommended that teachers select apps such as Appear, Geovin, Globe Observer, iNaturalist, PlantNet, and Seek since they have the greatest potential for taxonomic identification and offer the possibility of closer integration with scientific projects related to biodiversity conservation.

Keywords: Information and communication technologies, citizen science, biodiversity conservation.

Introduction

Although there are different perspectives on citizen science (CS), it commonly refers to a scientific program, led by professional scientists, where non-professional volunteers collect and/or analyse data which is subsequently utilised to advance scientific knowledge (MacPhail & Colla, 2020). These initiatives' growth has been enabled by the rise of technologies such as the internet, the web, and smartphones, which have made possible the development of apps that offer online recordings, real-time mapping, and photography (Mazumdar et al., 2018).

Moreover, technology provides an excellent opportunity for CS projects as it can

enhance accessibility for broader and more varied populations, including people who do not necessarily have background knowledge about the topics being studied, but who are willing to contribute by collecting data (Dehnen-Schumtz et al., 2016). Additionally, Web 2.0 tools have the potential to be powerful in implementing collective actions on environmental issues, exerting a significant influence on citizen empowerment (Reis, 2021).

Nowadays, we are experiencing a period of increasing environmental degradation and disconnect from nature in which public engagement and nature-based experiences are key (MacPhail & Colla, 2020). In this way, CS can promote learning about the value of biodiversity, stimulate and enrich decision-making, and commit to sustainable attitudes (Hadjichambis et al., 2020).

The potential of CS for non-formal and informal learning is well known. However, there is a limited number of studies about the connection of CS to formal education and its potential contributions within that context (Shah & Martinez, 2016). Many projects assess participants' enjoyment, but few explicitly assess their learning or changes in their attitudes toward science (Vitone et al., 2016). In this sense, it is possible that educators could use CS applications to promote their students' interest in a wide variety of scientific fields.

Therefore, we decided to delineate the characteristics of CS apps focused on biodiversity conservation to be employed as an alternative didactic strategy in pre-service and in-service teacher education in Córdoba, Argentina.

Method

Twenty free-download citizen science apps related to biodiversity were selected for the analysis. We considered their compatibility with the Android operating system—predominantly used in Argentina. Also, a methodology based on content analysis (Bardin, 1986) was used, a qualitative analysis tool was developed, and four dimensions were defined: multimedia, content, citizen engagement, and collaborative learning. Each dimension included analysis categories developed based on the proposal made by Finquelievich and Fischnaller (2014), and Garcia-Romano and Occelli (2019).

For the analysis of the 'multimedia' dimension, we considered the presence of text, photographs, images, and the possibility of accessing other resources such as links, data tables, and distribution maps. The 'content' dimension analysed the main biodiversity topics presented by the app.

The 'citizen engagement' dimension included five levels: (1) low level: volunteers only provide the capabilities of their mobile device for data collection, they do not necessarily know the project they are 'participating' in; (2) medium level: volunteers interact with ICT tools in the collection of data that will be analysed by researchers; (3) high level: volunteers collaborate in the collection

and monitoring of data, which will be analysed by researchers—they can join as volunteers at a research centre or in a scientific expedition; (4) advanced level: volunteers participate in the full extent of the scientific process, collaborating in the analysis of collected data and designing recording tools, potentially contributing to the definition of research objectives or hypotheses; (5) public policy level: citizens work alongside researchers in the processes of defining public policies that have technical or scientific components in the framework of a democratic process.

The 'collaborative learning' dimension included three levels: (1) low level: the app focuses on observation and data collection/submission, or only functions as an observation guide—some include participation in projects; (2) medium level: the app includes comments from other users and participation in projects, but does not include joint work with peers or specialists; (3) high level: the app involves working together with peers or specialists, and it may also include comments or private messages from other users, and participation in projects and missions.

Results

Concerning 'multimedia' characteristics, most apps (90%) contain photographs, texts with information for the user, and distribution maps (Figure 1a); within such a percentage, 50% of them included audio, video, and access to other links. Considering the 'content' dimension, it was found that, in general, apps combined multiple biodiversity contents in their proposal (Figure 1b), and only a couple of them focused on specific concepts (Figure 1c). Also, the study revealed that 50% of them included the description of the species and their complete taxonomy (Figure 1d).

As regards the 'citizen engagement' dimension, 30% of the apps scored level 1; 45%, level 2; and 25%, level 3; while levels 4 and 5 were not represented in the apps analysed (Figure 2). Concerning the dimension 'collaborative learning', 35% scored level 1; 35%, level 2; and 30%, level 3 (Figure 3).

Figure 1 a-b): Examples of the multimedia characteristics of different apps. Images taken from the (a-b) iNaturalist app; (c) Geovin app; (d) Seek app.

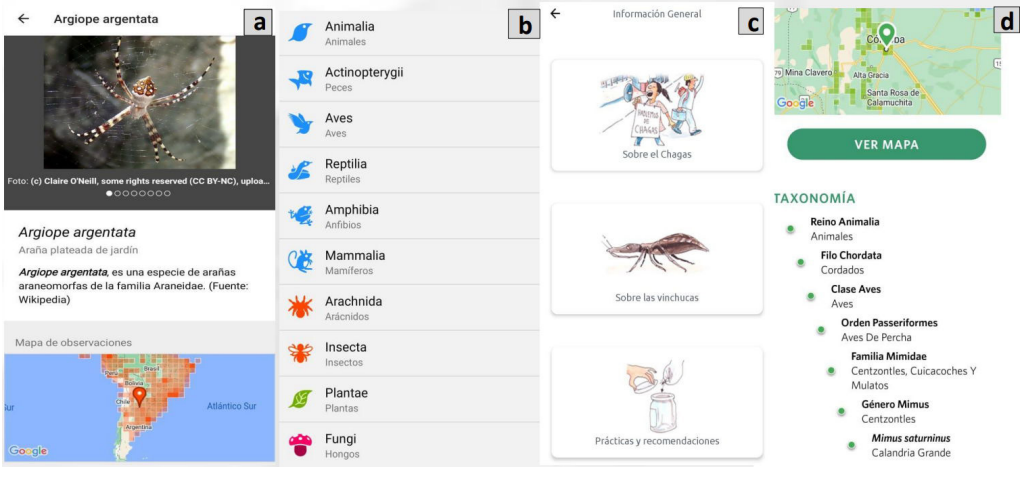


Figure 2. Percentage of apps in each level of the 'citizen engagement' dimension.

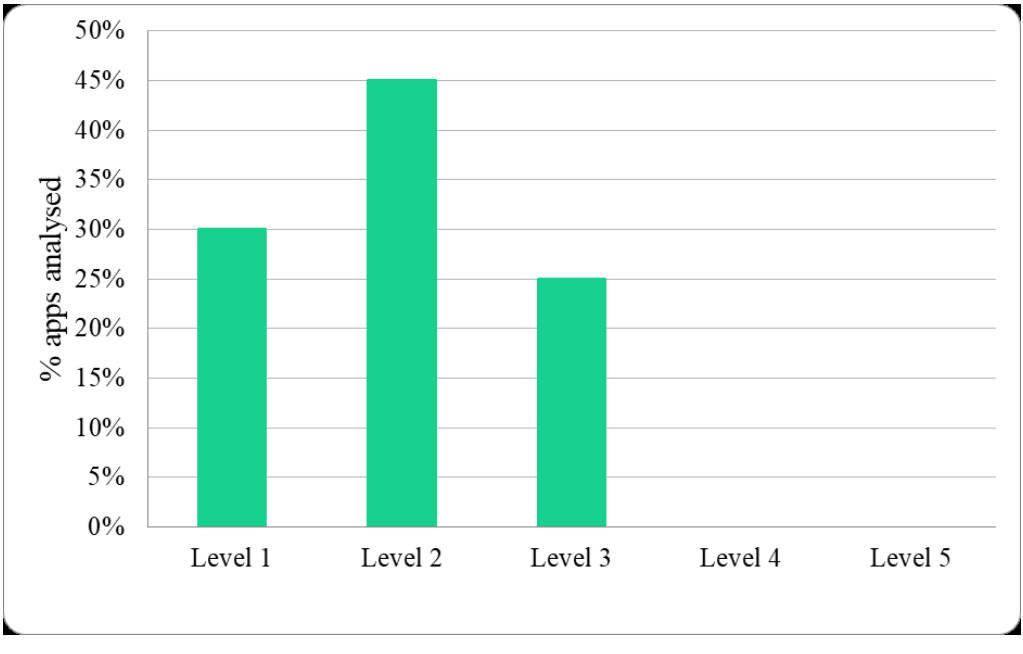
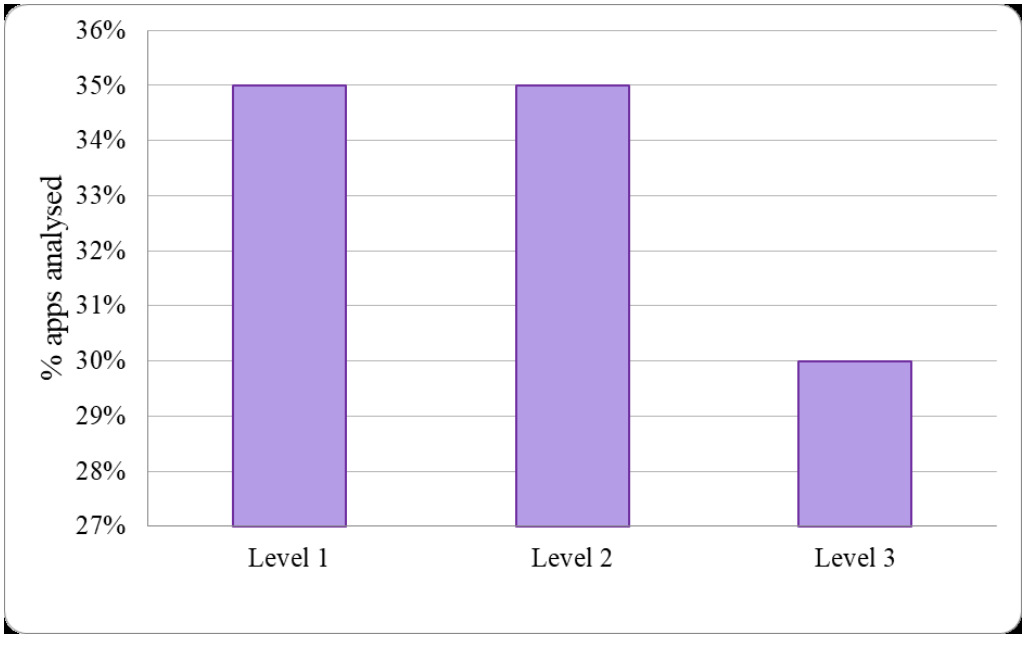


Figure 3. Percentage of apps in each level of the 'collaborative learning' dimension.



Discussion and Conclusions

In terms of 'multimedia' characteristics, most of the apps include photographs, texts with user information, and distribution maps/observations. The results show what was defined by Unger et al. (2020) regarding the potential of these apps to achieve a successful identification of organisms and obtain information on different taxonomic groups. Nevertheless, as highlighted by these authors, there is significance in instructing identification skills that integrate organism morphology alongside species recognition. One drawback of solely relying on a visual automated identification application is that students might overlook traditional morphological identification abilities essential for the precise application of traditional dichotomous keys (Unger et al., 2020).

Also, as outlined by Fundación Ciencia Ciudadana (2018), the 'contents' about biodiversity that can be addressed by the apps are multiple, requiring users to have different levels of background knowledge, thus allowing the development of a wide range of projects related to environmental education. As regards the 'citizen engagement' and the 'collaborative learning' dimensions, it has been found that several apps encourage volunteers to collaborate in the collection and monitoring of data that will be analysed by researchers, giving citizens the possibility to participate in scientific projects. Although many apps promote

collaboration between volunteer citizens and specialists, many apps only offer volunteers the possibility of collecting data. These results coincide with those presented by Finquelievich and Fischnaller (2014) who point out that citizen science is not always an 'open science'.

This work is aimed at (re)thinking how CS apps could be integrated into pre-service and in-service education in Córdoba, Argentina. Based on this, the apps *Appear, Geovin, Globe Observer, iNaturalist, PlantNet, and Seek* had the greatest potential for the taxonomic identification of nearby organisms and also presented the possibility of stronger integration with scientific projects related to biodiversity conservation. The characteristics of these apps are expressed in Table 1.

Table 1. Characteristics of the apps with the greatest potential.

Mobile application	Multimedia	Content	Citizen Engagement	Collaborative Learning
Appear	Texts, photographs, links, distribution maps	Ecology of rivers, streams, lakes, and estuaries	(3) High Level	(2) Medium Level
Geovin	Photographs, distribution maps	Distribution of the triatomines in Argentina	(2) Medium Level	(3) High Level
Globe Observer	Texts, photographs, distribution maps	It includes four sub-applications: Clouds, Mosquito Habitat Mapper, Land Cover, and Trees	(2) Medium Level	(2) Medium Level
iNaturalist	Texts, photographs, links, data tables, distribution maps, audio	Different taxa Genus/family/ species description	(3) High Level	(3) High Level
PlantNet	Texts, photographs, links, distribution maps, audio	Wild plants (useful, cultivated, and ornamental)	(3) High Level	(3) High Level
Seek	Texts, photographs, links, data tables, distribution maps, audio	Different taxa Endemic, native, and introduced species State of conservation	(2) Medium Level	(3) High Level

Acknowledgement

This study was funded by the National Interuniversity Council (CIN, Argentina), the National Scientific and Technical Research Council (CONICET) [PIBAA 28720210101027CO], Desafío IMPACTAR 115 (Ministry of Science, Technology and Innovation, Argentina) and by the Science and Technology Teaching Department of the National University of Córdoba (SECyT-UNC) [336 202301 00484 CB].

References

Bardin, L. (1986). *El análisis de contenido*. Akal.

Dehnen-Schmutz, K., Foster, G. L., Owen, L., & Persello, S. (2016). Exploring the role of smartphone technology for citizen science in agriculture. *Agronomy for Sustainable Development*, 36, 1-8. https://doi.org/10.1007/s13593-016-0359-9

Finquelievich, S., & Fischnaller, C. (2014). Ciencia ciudadana en la Sociedad de la Información: nuevas tendencias a nivel mundial. *Revista Iberoamericana de Ciencia Tecnología y Sociedad*, 9(27), 11-31.

Fundación Ciencia Ciudadana. (2018). *Ciencia ciudadana: Principios herramientas proyectos de Medio Ambiente*. Embajada de Canadá.

Garcia Romano, L., & Occelli, M. (2019). Un modelo analítico para caracterizar recursos tecnológicos basados en contenidos científicos. *Revista de Enseñanza de la Física*, 31(1), 15–25. https://revistas.unc.edu.ar/index.php/revistaEF/article/view/24667

Hadjichambis, A.Ch., Reis, P., Paraskeva-Hadjichambi, D., Činčera, J., Boeve-de Pauw, J., Gericke, N., & Knippels, M.C. (2020). *Conceptualizing Environmental Citizenship for 21st Century Education*. Springer Open.

MacPhail, V.J., & Colla, S.R (2020). Power of the people: A review of citizen science programs for conservation. *Biological Conservation*, 249, 1-15. https://doi.org/10.1016/j.biocon.2020.108739

Mazumdar, S., Ceccaroni, L. Piera, J., Hölker, F., Berre, A., Arlinghaus, R., & Bowser, A. (2018). Citizen science and technologies and new opportunities for participation. In S. Hecker, M. Haklay, A. Bowser, Z. Makuch, J. Vogel, & A. Bonn (Eds.), *Citizen Science - Innovation in Open Science, Society and Policy*. (pp. 303-320). UCL Press.

Reis, P. (2021). Challenges to Science Education in Troubled Times. *Ciência & Educação*, 27. https://doi.org/10.1590/1516-731320210000

Shah, H. R., & Martinez, L. R. (2016). Current Approaches in Implementing Citizen Science in the Classroom. *Journal of Microbiology & Biology Education,* 17(1), 17-22. https://doi.org/10.1128/jmbe.v17i1.1032

Unger, S., Rollins, M., Tietz, A., & Dumais, H. (2020). iNaturalist as an engaging tool for identifying organisms in outdoor activities. *Journal of Biological Education,* 55(5), 537-547. https://doi.org/10.1080/00219266.2020.1739114

Vitone, T., Stofer, K., Steininger, M. S., Hulcr, J., Dunn, R., & Lucky, A. (2016). School of Ants Goes to College: Integrating Citizen Science into the General Education Classroom Increases Engagement with Science. *Journal of Science Communication, 15(1), A03.* https://doi.org/10.22323/2.15010203

The Effect of the Instructional Technologies Course Taught with The Blended Learning Model on the Pre-Service Teachers' Self-Efficacy Beliefs in Web 2.0 Practical Content Development

Betül Karaduman[1] *, Pınar Fettahlıoğlu*[2]

[1]Çukurova University, Türkiye

[2]Çukurova University, Türkiye

The aim of this study is to determine the effect of the instructional technologies course taught with the blended learning model on the pre-service teachers' self-efficacy beliefs in developing content quickly. The study group consists of 44 science teacher candidates (21-experimental group, 23-control group) who attend the science teaching program and take the instructional technologies course. A mixed method was used as a research design; particularly, a special type of Concurrent Nested Strategy was applied. The quantitative portion was based on the a quasi-experimental design with unequal pretest-posttest control group from the quantitative research, and the qualitative portion was based on the holistic multiple case study method. The quantitative portion of the research was conducted with 21 (experimental group), 23 (control group) second-year pre-service science teachers studying at a state university. The qualitative portion of the study was conducted with six pre-service science teachers selected among the 21 (experimental group) pre-service science teachers based on the pre-test results obtained from quick content development self-efficacy belief scale. As quantitative data collection tools, the web 2.0 practical content development self-efficıency scale developed by Birişçi et al. (2018) were used. As a qualitative data collection tool, a semi-structured opinion form developed by the researchers was used. t tests for dependent groups and independent sample t-test were used to analyze quantitative data. Both descriptive and content analyses of the qualitative data were performed. In the analysis of the data, it was determined that only the students in the experimental group developed web 2.0 practical content development self-efficacy belıef.

Keywords: Instructional technologies, self-efficacy belief, blended learning model.

Introduction

Rapid changes and advances occur in education and training technologies as a result of technology being an essential part of our lives. Teachers must have

the competences and skills to use these technologies concurrently in order to integrate these advancements into the classroom setting. Given that the faculties of education, where pre-service teachers receive their training, are at the forefront of the places where teachers acquire knowledge and skills, general information about instructional technologies should be provided to pre-service teachers in these areas, along with identifying the skills necessary for them to use these technologies and providing the necessary environment for these skills to be acquired. When the pertinent literature is evaluated, it becomes clear that self-efficacy belief is a quality that instructors should possess. This belief is generally expressed by Bandura (1982) as "personal judgments about what to do in possible situations that people may encounter". In other words, self-efficacy can also be expressed as a person's belief in organizing and applying the necessary skills in order to bring about targeted and desired results (Bandura, 1997). Accordingly, self-efficacy belief in individuals shows the belief that one has in order to perform that skill rather than the ability to perform a skill. Therefore, it is conceivable to accept teachers' self-efficacy views as their beliefs about incorporating technology opportunities within the context of educational activities provided their beliefs are assessed in the context of evolving technologies (Chen, 2008). Teachers' technology self-efficacy is related to their ability to perform certain tasks as a result of their use of Web 2.0 tools. While Pan and Franklin (2011) emphasize that teachers' self-efficacy is an important factor in the inclusion of Web 2.0 technology, they state that if a teacher does not trust their abilities, it is likely that they will fail to include Web 2.0 tools in their lessons. In this situation, it's critical to offer appropriate settings for teacher candidates to develop this ability. In this context, one of the models that has been emphasized recently is the blended learning model.

According to Smith and Hill (2019), blended learning encompasses all technically supported learning environments except online-only learning environments and classroom teaching. According to So and Bonk (2010), some questions need to be answered in order to determine the "right mix" between face-to-face classroom teaching and online learning. The first of these questions is, what part of the interaction should take place in classroom-based face-to-face or online settings? Second, when should online or classroom-based face-to-face learning be used? Finally, how can the two methods be combined to optimize the learning process? Generally speaking, the blended learning model integrates traditional face-to-face classroom teaching with online digital learning. In addition, it is stated that the harmony and fusion of the online learning environment and the face-to-face learning environment is important for the realization of learning (Sungur Alhan, 2020). In this context, the aim of this study is to determine the effect of the instructional technologies course, which is taught with the blended learning model, on the web 2.0 practical content development self-efficacy beliefs of teacher candidates.

Method

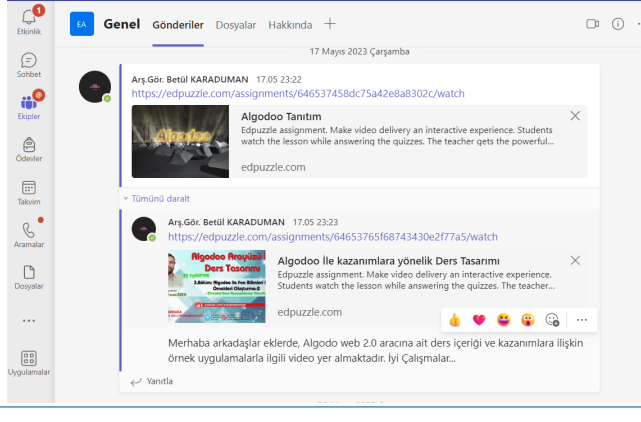
A mixed method was used as a research design; particularly, a special type of Concurrent Nested Strategy was applied. The quantitative portion was based on a quasi-experimental design with unequal pretest-posttest control group from the quantitative research, and the qualitative portion was based on the holistic multiple case study method (Creswell and Plano-Clark 2011).

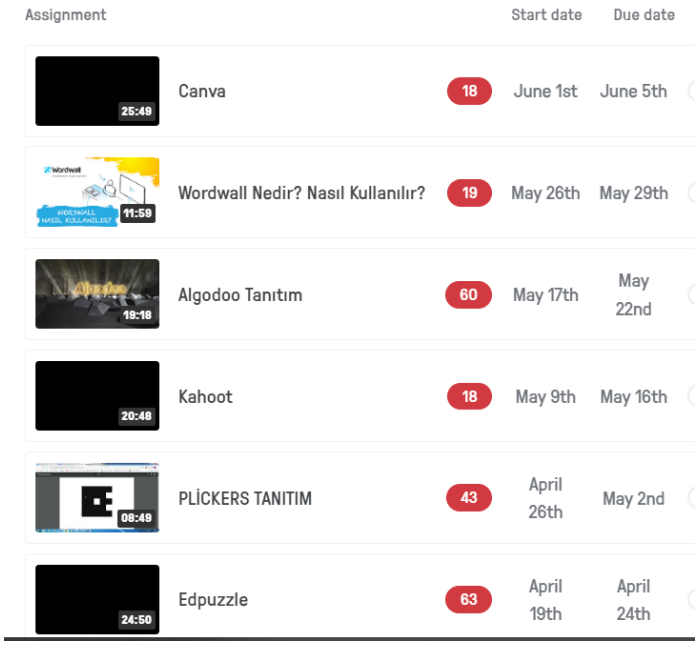
Procedure:

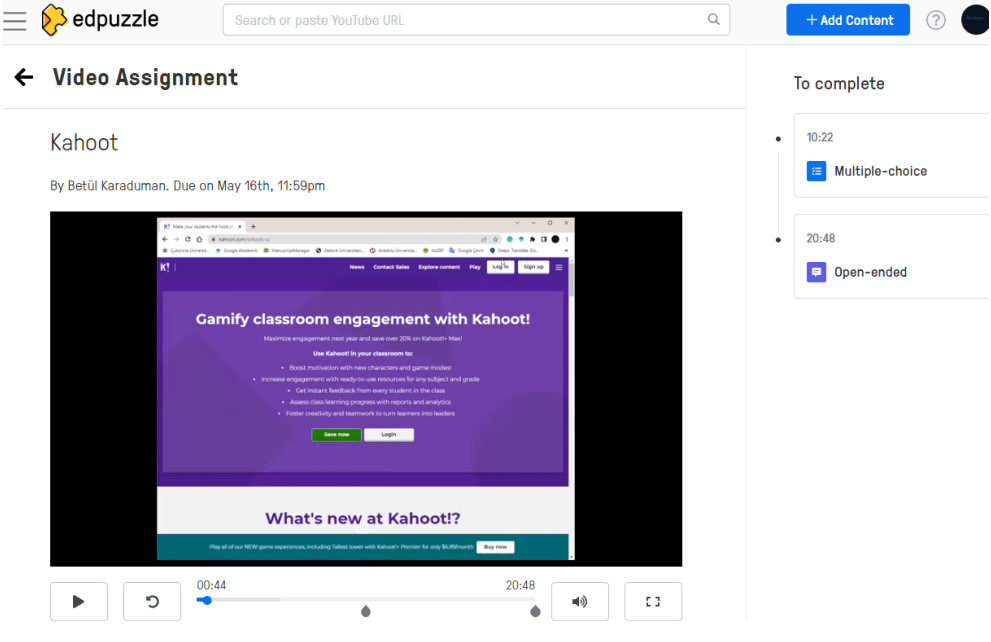
- In the first four weeks of the course, the necessary basic information was presented by the instructor, starting from the 5th week, the students were divided into groups. Each week, the groups came to the lesson by making preparations for the web 2.0 tool of that week through the materials sent to them via Edmodo system. In the lesson, they selected any learning outcome in the science curriculum and prepared a lesson plan for that learning outcome according to the 5E learning model, including the related web 2.0 tool, and the web 2.0 tool appropriate to this lesson plan. Web 2.0 tools included in the lesson plans created according to the 5E learning model are Eddpuzzle, Algodo, Plickers, Wordwall, Kahoot and Canva.
- The prepared lesson plans were uploaded to the Microsoft teams LMS system and evaluated by the instructor every week.
- Focus group interviews were conducted with selected students every week.
- At the end of 8 weeks in which the flipped learning model was applied, the scales were reapplied as a post-test and interviews were conducted with the students again.

You can see the screenshots of the process of the flipped learning process in Figure 1.

Figure 1: Screenshots of the implementation of the flipped learning process







Participants

The quantitative portion of the research was conducted with 21 (experimental group), 23 (control group) second-year pre-service science teachers studying at a state university. The qualitative portion of the study was conducted with six pre-service science teachers selected among the 21 (experimental group) pre-service science teachers based on the pre-test results obtained from quick content creation self-efficacy belief scale.

Data Collection Tools

As quantitative data collection tools, the web 2.0 practical content development self-efficiency scale developed by Birişçi et al. (2018) were used. As a qualitative data collection tool, a semi-structured opinion form developed by the researchers was used.

Web 2.0 practical content development self-efficiency scale

The scale, which consists of 21 items, was prepared in a 5-point Likert type and its categories range from I am very inadequate to I am very sufficient. The factors of the scale gathered under 3 factors were designed as the usability of Web 2.0 tools in the process of preparing the course content, presenting the course content and evaluating the learning outcomes. The Cronbach - Alpha internal consistency coefficient of the scale was found as $\alpha = 0.955$. For this study, it was calculated as .90.

A Semi-Structured Opinion Form

Interviews that were conducted to obtain information about the process were completed at the end of each class and with each individual participant. When preparing the questions for the semi structured interviews, the following steps were taken (Cresswell and Plano-Clark 2011):

1. Before the interview questions were formed, the general and special purpose of the research was listed in draft form.
2. The questions were prepared in such a way that they would not cause the pre-service science teachers to be bored or ashamed and so that the questions would be easy to answer using clear explanations.
3. The questions were evaluated by considering the perspectives of experts in the field.
4. A pilot study was conducted of the developed questions. Ten pre-service science teachers teaching second grade were interviewed for the pilot study. The pilot interviews were conducted during the pilot implementation stage of the study. The results of the implementation were satisfactory and reviewed by an expert before the interview questions were refined into the final form

Data Analysis

Independent samples t-test was used to analyze quantitative data. Both descriptive and content analyses of the qualitative data were performed.

Results

Pre-service teachers' self-efficacy beliefs for web 2.0 practical content development

Independent groups t-test was performed to determine whether there was a change in the self-efficacy beliefs of the pre-service teachers in the experimental group where the flipped learning model was applied and in the control group where the traditional learning model was applied. The results are presented in Table 1.

Table 1. ***Independent samples t-test results on pre-service teachers' self-efficacy beliefs for developing practical content***

		N		S	t	sd (df)	p (sig)
Preparation	Experimental group	21	4,3736	,42223	3,359	42	0.02
	Control group	23	3,8495	,59016	3,409		
Presentation	Experimental group	21	4,3690	,50385	1,662	42	.104
	Control group	23	4,0870	,61055	1,677		
Evaluation	Experimental group	21	4,6310	,45152	2,620	42	.012
	Control group	23	4,2174	,58048	2,650		
Overall average	Experimental group	21	4.42	.42	3.006	42	.004
	Control group	23	3.97	.57			

According to Table 1, there is a significant difference in pre-service teachers' self-efficacy beliefs of developing practical content according to whether the Instructional Technologies course is taught with the flipped learning model or not [$t_{(42)}$= 3.006, p<.005]. Practical content development self-efficacy beliefs of the experimental group in which flipped learning was applied (=4.42) were higher than the self-efficacy beliefs of the control group in which the traditional model was applied (=3.97). When evaluated in terms of the sub-dimensions of the scale, it is seen that there is a statistically significant difference in terms of the self-efficacy of pre-service teachers in the preparation and evaluation sub-dimensions, except for the presentation sub-dimension. In the presentation sub-dimension, although the mean of the experimental group was higher, it was not statistically significant. When this finding is evaluated in general, it can be interpreted that teaching the instructional technology course according to the flipped learning model has a positive effect on the self-efficacy beliefs of pre-service teachers for practical content development. However, it is thought that

the reason for the difference in the presentation sub-dimension is that a web 2.0 tool for blogging was not included during the course but it was included in the scale (Due to the earthquake disaster in our country, the course content had to be changed.).

The themes formed according to the results of the content analysis of the interviews conducted with pre-service teachers are presented in Table 2.

Table 2. The themes that emerged as a result of the interviews with pre-service teachers.

Sub-dimensions	Themes	Basic concepts and design models	The relationship between learning theories and instructional technology	5E learning model	Class dojo	Edpuzzle	Plickers	Kahoot	Algodo	Wordwall	Canva	Pawtoon
Preparation	Create a worksheet		S5, 6	S3,4 S5, 6	S3,4 S5, 6	S1,2; **S3,4** S5, 6	**S1,2**; **S3,4** S5, 6	**S1,2**; **S3,4** S5, 6	**S1,2**; **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6
	Create videos, animations		S5, 6	S5, 6	S5, 6	S3,4 S5, 6	S3,4 S5, 6	**S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6
	Guidance in using Web 2 tools		S5, 6	S5, 6	S5, 6	S3,4 S5, 6	S3,4 S5, 6	S3,4 S5, 6	S3,4 S5, 6	S1,2 S3,4 S5, 6	S1,2 S3,4 S5, 6	S1,2 S3,4 S5, 6
	Using Web 2 tools in a pedagogical way		S3,4 S5, 6	S1,2 S3,4 S5, 6	S1,2 S3,4 S5, 6	S1,2 **S3,4** S5, 6	S1,2 **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6
Presentation	Photo, video sharing	**S3,4** S5, 6	**S3,4** S5, 6	S1,2 **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6
	Blog sharing	S5, 6	S5, 6	S5, 6	S5, 6	S5, 6	**S3,4** S5, 6	**S3,4** S5, 6	S2,1 **S3,4** S5, 6	S2,1 **S3,4** S5, 6	S2,1 **S3,4** S5, 6	S2,1 **S3,4** S5, 6
	Presentation sharing	**S3,4** S5, 6	**S3,4** S5, 6	S1,2 **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6

Assessment	Interactive questions		S5, 6	**S3,4** S5, 6	**S3,4** S5, 6	S1,2 **S3,4** S5, 6	S1,2 **S3,4** S5, 6	S1,2 **S3,4** S5, 6	S1,2 **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6
	Test question		S5, 6	**S3,4** S5, 6	**S3,4** S5, 6	S1,2 **S3,4** S5, 6	S1,2 **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6
	Puzzle		S5, 6	**S3,4** S5, 6	**S3,4** S5, 6	S1,2 **S3,4** S5, 6	S1,2 **S3,4** S5, 6	S1,2 **S3,4** S5, 6	S1,2 **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6
	Preparing different measuring instruments		S5, 6	**S3,4** S5, 6	**S3,4** S5, 6	**S3,4** S5, 6	S1,2 **S3,4** S5, 6	S1,2 **S3,4** S5, 6	S1,2 **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6	**S1,2** **S3,4** S5, 6

According to Table 2; in general, it is seen that the development in self-efficacy beliefs was realised after the 5th week, that is, after they started the practice. Also, It is seen that the effect on the self-efficacy of pre-service teachers with low self-efficacy was realized when the course started to be taught with flipped learning. It is seen that the change in the self-efficacy belief of the middle group started from the first week, but this change was limited to the presentation sub-dimension. It is thought that the reason for this is that they were familiar with web 2.0 tools related to presentation before. Some quotes from these teacher candidates are as follows;

> "I knew web 2 tools before, but I did not know at which stage of the course to use them, but after this course we took in 2nd grade, I can say that I started to use them consciously and tried to improve them. This course makes me feel more equipped than before(S6)."

> "Preparing presentations from these applications increased my self-confidence. I always thought that such applications would be used by those who understand computers, and now I believe that I can do it too. Since we started to flipped classroom from the 4th week, I started to improve (S2)."

> "I think I am competent enough to apply the programs in general. The general process of the course was very effective in gaining this competence (S4)."

It is seen that the pre-service teachers in the upper group reported that their already existing self-efficacy had increased even more since the second week. Some quotes from these teacher candidates are as follows;

> "When we first started the course, I only had an idea about kahoot and I felt inadequate and I had no idea how to do it. Later, as I started to use web.2.0 tools and apply them to lesson plans, I started to feel more competent. Especially towards the last weeks, I started to do it easily. But

especially after Edpuzzle, I got better(S1)."

"The course has provided me with very valuable useful information, I was using web 2 tools before this course, but I think I have learned more diverse tools with this course, I will use web 2.0 tools frequently in different subjects in my professional life (S5)."

Conclusion and Discussion

The study's results revealed that the experimental group, which received instruction using a blended learning paradigm, had mean self-efficacy scores that were considerably higher than those of the control group. This result suggests that the use of the flipped classroom model, one of the blended learning models, in the instructional technologies course had a favorable impact on pre-service teachers' judgments of their own self-efficacy in developing practical web 2.0 content. When the relevant research is evaluated, it becomes clear that comparable conclusions have been drawn about how this model affects students' affective traits like attitude, motivation, and self-efficacy (Gürsoy,2018; Ünlütürk, 2022; Debbağ, 2018; Souza & Rodrigues, 2015; Kenna, 2014).

The flipped classroom model is a method of instruction that gives students the opportunity to learn effectively by giving them the feedback they require as they participate in instruction in both the classroom and outside of the classroom while utilizing modern technology (Aydın, 2016). Since the parts that are not understood or missing during the work at home are the parts that the student needs the teacher the most, their reinforcement is saved for the classroom environment (Stone, 2012). Students get more time to work with peers, engage with the material more deeply, and receive immediate feedback from their teacher. (Hamden et al., 2013).

References

Aydin, B. (2016). Ters yüz sınıf modelinin akademik başarı, ödev/görev stres düzeyi ve öğrenme transferi üzerindeki etkisi [Unpublished master's thesis]. University of Süleyman Demirel.

Bandura, A. (1982). Self-efficacy mechanism in human agency. American Psychologist, 37(2), 122-147.

Bandura, A. (1997). Self-efficacy: The exercise of control. New York: Freeman.

Birişçi, Kul, Aksu, Akaslan, & Çelik (2018). Web 2.0 hizli içerik geliştirme öz-yeterlik algısını belirlemeye yönelik ölçek (W2ÖYAÖ) geliştirme çalışması. *Eğitim Teknolojisi Kuram ve Uygulama*, *8*(1), 187-208.

Chen, Y. (2008). A mixed-method study of EFL teachers' internet use in language instruction. Teaching and Teacher Education, 24, 1015-1028

Creswell, J. W., & Plano-Clark, V. L. (2011). Designing and conducting mixed methods

research (2nd ed.). Thousand Oaks, CA: Sage

Debbağ, M. (2018). Öğretim ilke ve yöntemleri dersi öğretim programı için hazırlanan ters-yüz edilmiş sınıf modelinin etkililiği [Unpublished doctoral thesis]. University of Bolu İzzet Baysal.

Gürsoy G. (2018). Fen Öğretiminde Okul Dışı Öğrenme Ortamları [Outdoor learning environments in science education]. *Turkish Studies,13*(11), 623-649. https://doi.org/10.7827/TurkishStudies.13225

Hamden, N., McKnight, P.E., McKnight, K., & Arfstrom, K. (2013). A review of flipped learning. Flipped Learning Network. Upper Saddle River, NJ: Pearson Education, Retrieved May 9, 2015, from: http://www.flippedlearning.org/cms/lib07/VA01923112/Centricity/Domain/41/LitReview_FlippedLearning.pdf

Kenna, D. C. (2014). *A study of the effect the flipped classroom model on student self-efficacy* (Doctoral dissertation, North Dakota State University).

Pan, S. C. & Franklin, T. (2011). In-Service teachers' self-efficacy, professional development, and web 2.0 tools for ıntegration. New Horizons in Education, 59(3), 28-40.

Smith, K., & Hill, J. (2019). Defining the nature of blended learning through its depiction in current research. Higher Education Research and Development, 38(2), 383–397. https://doi.org/10.1080/07294360.2018.1517732

So, H.-J., & Bonk, C. J. (2010). Examining the roles of blended learning approaches in computer-supported collaborative learning (CSCL) environments: A Delphi study. Journal of Educational Technology & Society, 13(3), 189–200

Souza, M., & Rodrigues, P. (2015). Investigating the effectiveness of the flipped classroom in an introductory programming course. *The New Educational Review*, *40*(1), 129-139. https://doi.org/10.15804/tner.2015.40.2.11

Stone, B. B. (2012). Flip your classroom to increase active learning and student engagement. In *Proceedings from 28th Annual Conference on Distance Teaching & Learning, Madison,* Wisconsin, U.S.(pp.1-5).

Sungur Alhan, S. (2020). The opinions of pre-service science teachers about the blended learning environment. Journal of Erzincan University Faculty of Education, 22(2), 397–414. https://doi.org/10.17556/erziefd.633189

Ünlütürk, A. Ö. (2022). Ters yüz öğrenme ile yapılandırılmış okul dışı fen eğitiminin çeşitli değişkenler bakımından incelenmesi. [Unpublished master›s thesis]. University of Karamanoğlu Mehmetbey.

Digital Literacy in Turkish High School Science Textbooks

Aliye Tuba Gülsoy[1], Mustafa Çakır2

[1]Ministry of National Education, Türkiye

[2]Marmara University, Türkiye

In the 21st century, many innovations have been reflected in the educational sector and changed teaching and learning practices. Changes in the process of acquiring and using knowledge, differentiating student characteristics, innovations in accessing digital resources, and changing understandings of what to learn and how to learn required rethinking traditional methods and reforming education. One of the skills that entered the curriculum and textbooks with the education reforms is digital literacy. In this context, it is necessary to examine science textbooks in terms of digital literacy. This study aims to examine the high school science textbooks approved and distributed by the Turkish Ministry of National Education in terms of digital literacy. Within the scope of the study, a textbook evaluation rubric was developed based on the UNESCO Digital Literacy Competence Framework. The study was carried out with a qualitative research approach and document analysis was used as a research design. Research data were analyzed by the comparative descriptive analysis method. The references to digital competence areas in textbooks increase with the increase in grade level. In the books examined, the area of competence that takes place at the highest rate at all grade levels is "information and data literacy". There is no content in the textbooks within the scope of "career-related competencies". The level of including "security" and "problem-solving" competencies in the reviewed books is quite low. It can be expected that references to "problem-solving" competencies would be higher in high school science textbooks since high school students are expected to have required higher-order thinking skills and required cognitive abilities. Additionally, it is very important to provide the new generation of students, who are digital natives, with security-related competencies such as understanding the risks and threats in the digital environment and protecting personal data and privacy.

Keywords: Digital literacy, science textbook, evaluation rubric.

Introduction

Changes in the process of acquiring and using information, differentiating student characteristics, innovations in online access, connecting and sharing, and changing understandings of what to learn and how to learn have made it necessary to rethink traditional methods and reform education in the 21st

century. From this point of view, which digital skills individuals should have and their reasons were questioned (OECD, 2005), and strategies were developed for the application of these skills in educational areas. In this context, many conceptual frameworks have been put forward, the main objective of which is to guide national curriculum policies for the implementation of 21st-century qualifications in education (Erstad & Voogt, 2018; Voogt & Roblin, 2010). Among these initiatives, P21, OECD, and EU presented a framework that provides the conceptualization of 21st-century core competencies. At the same time, ATCS, NAEP focused on assessment, ISTE, En Gauge, and UNESCO focused on digital literacy and integration of technology into the curriculum (Voogt & Roblin, 2010). Although they focus on different areas, collaboration, communication, and information and communication technologies (ICT) literacy have been included in all of these frameworks, and many curricula have been organized around these skills (Voogt & Roblin, 2010). With the digitalization of all areas of life, a new skill area presented as "digital competence" or "digital literacy" has taken its place in the education programs of different countries (Erstad, 2013; Erstad & Voogt, 2018).

What is Digital Literacy?

Digital literacy, in short, is the knowledge, skills, and attitudes that enable us to cope with all the complexities that emerge with the changing world order and to participate in the digital society (Feerrar et al., 2019). The European Commission defines digital literacy as follows.

> "Digital literacy is the ability to access, manage, understand, integrate, communicate, evaluate, create, and disseminate information safely and appropriately through digital technologies. It includes competencies that are variously referred to as information literacy and media literacy, computer, and ICT literacy. Digital Literacy involves a dimension of active and civic engagement with the digital world and promotes active citizenship" (European Commission, 2022, p. 11).

According to this definition, digital literacy covers a broad meaning from acquiring knowledge to participatory citizenship and is associated with a wide range of knowledge, skills, and competencies. It is important to provide all young people, without discrimination, with these skills, knowledge, and understanding that will enable them to recognize and question the various effects of digital technologies while taking advantage of the opportunities of the digital age (Payton & Hague, 2010). In this context, educators and researchers have determined the theoretical framework and competencies of digital literacy and called for curriculum design around these competencies (Neumann et al., 2017; Osterman, 2013). Addressing digital literacy throughout the curriculum will provide young people with competencies that will support their full participation

in society, both now and in the future, and their ability to make the most of developing digital technologies and social-cultural practices (Neumann et al., 2017; Osterman, 2013).

In light of developments in the world, Turkey is one of the countries that has updated its education program by considering the requirements of the 21st century. The 2018 curriculum, which was developed by making use of the European Qualifications Framework, was aimed at raising digitally literate individuals, and the relevant standards were integrated into the curriculum.

Textbooks play a critical role in achieving the goals of raising digitally literate individuals. Textbooks maintain their importance as a component of teaching by improving teaching at all levels of schools and in different disciplines, guiding the scope and content of the course, and being a guide of the course (Knight, 2015; Knight & Horsley, 2011). Due to its active role in education and training, digital literacy skills should be supported with textbooks along with other competencies. Due to its nature that concerns all disciplines, digital literacy should not be limited to technology-related parts of a particular course or curriculum but should be included in all textbooks. When textbooks from different disciplines are examined in the context of digital literacy, no study examines Turkish high school science textbooks based on global digital literacy competence areas. For this reason, it was necessary to examine the availability of digital literacy skills in high school science textbooks. In this context, the research aims to examine the high school science textbooks used in schools affiliated with the Ministry of National Education of Turkey in terms of digital literacy.

Method

One of the qualitative research approaches is document analysis. In this study, interpretive document analysis was employed to examine science textbooks with all their contents. The research data were collected from the 9th, 10th, 11th, and 12th-grade physics, chemistry, and biology textbooks taught in schools affiliated with the Ministry of National Education of Turkey in the 2022-2023 academic year. Textbooks were provided as printed copies as they were distributed in schools, and PDF (Portable Document Format) formats were used when necessary. In the data collection process, an evaluation chart was needed to analyze which and at what level the digital literacy competence areas the textbooks contain and to ensure the reliability of the research. A content analysis chart was developed in line with the proposed UNESCO digital literacy competencies framework by making additions to DigComp 2.1, and the books were analyzed with this tool. UNESCO “A Global Framework of Reference for Digital Literacy Skills for Indicator 4.4.2” digital literacy competence areas are “Fundamentals of hardware and software”, “Information and data literacy”,

"Communication and collaboration", "Digital content creation", "Safety", "Problem solving", "Career-related competences". Digital literacy competency areas and competencies proposed by UNESCO are presented in Table 1.

Table 1. Proposed digital literacy competence areas and competencies (Antoninis ve Montoya, 2018)

Competence area	Competences
0. Devices and software operations	0.1 Basic knowledge of hardware such as turning on/ off and charging, locking devices 0.2 Basic knowledge of software such as user account and password management, login, and how to do privacy settings, etc.
1. Information and data literacy	1.1 Browsing, searching and filtering data, information and digital content 1.2 Evaluating data, information and digital content 1.3 Managing data, information and digital content
2. Communication and collaboration	2.1 Interacting through digital technologies 2.2 Sharing through digital technologies 2.3 Engaging in citizenship through digital technologies 2.4 Collaborating through digital technologies 2.5 Netiquette 2.6 Managing digital identity
3. Digital content creation	3.1 Developing digital content 3.2 Integrating and re-elaborating digital content 3.3 Copyright and licenses 3.4 Programming
4. Safety	4.1 Protecting devices 4.2 Protecting personal data and privacy 4.3 Protecting health and well-being 4.4 Protecting the environment
5. Problem solving	5.1 Solving technical problems 5.2 Identifying needs and technological responses 5.3 Creatively using digital technologies 5.4 Identifying digital competence gaps 5.5 Computational thinking
6. Career-related competences	6. Career-related competencies refers to the knowledge and skills required to operate specialized hardware/software for a particular field, such as engineering design software and hardware tools, or the use of learning management systems to deliver fully online or blended courses.

Note. Underscored competence areas and competencies are proposed additions to the existing DigComp 2.1 competencies.

The items of the analysis rubric were reviewed by the experts and the rubric was revised to its final form after the necessary changes were made according to experts' opinions. To determine the reliability of the data analysis rubric, the

11th-grade chemistry textbook was analyzed by two experts using the same rubric. The reliability coefficient was calculated with the Holsti equation, which calculates the percentage of agreement between the decisions of at least two coders to encode the same data unit (0.91). In qualitative data analysis, it is very important to organize the data with methodical rigor and effectiveness (O'leary, 2017). The data were described in the study according to UNESCO digital literacy themes and the results were interpreted and presented. The steps followed in the data analysis process of the study were as follows: identifying the contents related to digital literacy, coding the contents found per the evaluation form criteria, calculating the percentages, and presenting the data in tables.

Findings and Results

When high school biology, physics, and chemistry textbooks were analyzed in terms of digital literacy, a total of 557 codes were made. A total of 5 competencies were included in biology textbooks, 4 competencies were included in physics textbooks, and 6 competencies were included in chemistry textbooks. The digital literacy competence area frequency table of high school science textbooks is presented in Table 2.

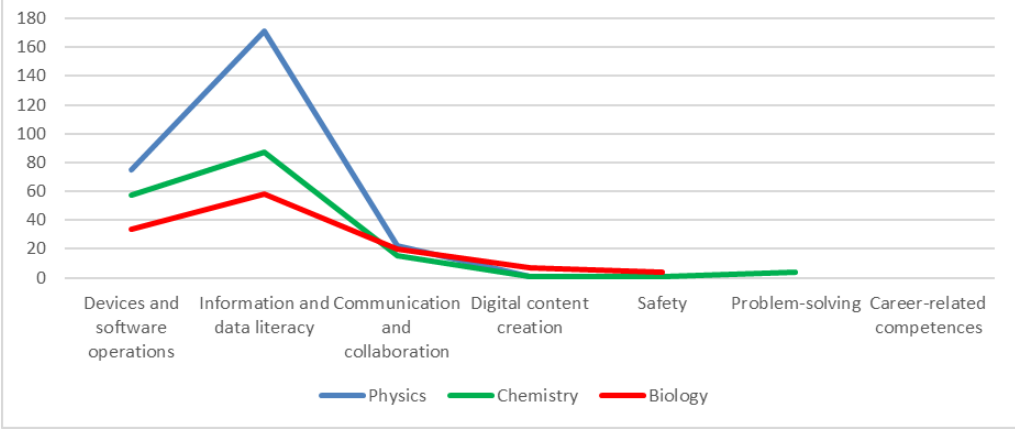
Table 2. Competence areas and their representations in Turkish physics, chemistry and biology textbooks

Competence area	Physics		Chemistry		Biology		Total	
	N	%	N	%	N	%	N	%
Devices and software operations	75	27.9	57	34.6	34	27.6	166	29.9
Information and data literacy	171	63.5	87	52.7	58	47.2	316	56.7
Communication and collaboration	22	8.2	15	9.1	20	16.3	57	10.2
Digital content creation	1	0.4	1	0.6	7	5.7	9	1.6
Safety			1	0.6	4	3.2	5	0.9
Problem-solving			4	2.4			4	0.7
Career-related competences								
Total	269	100	165	100	123	100	557	100

Table 2 shows that references to the competency areas of "devices and software operations", "information and data literacy", "communication and collaboration" and "digital content creation" are found in biology, physics, and chemistry textbooks. The "information and data literacy" competence area is the most frequently coded domain for all three courses with 47.2% in biology

textbooks, 63.5% in physics textbooks, and 52.7% in chemistry textbooks. In chemistry textbooks, there was one code each for "security" and "digital content creation". In physics textbooks, "safety" and "problem-solving" competency areas were not coded. Biology textbooks also did not include competencies related to "problem-solving". None of the science textbooks were coded for "career-related competencies". The distribution of the number of references to digital competency areas in the textbooks according to the courses is given in Figure 1.

Figure 1. Distribution of competence areas in physics, chemistry, and biology textbooks



The graph above shows that physics textbooks contain the highest number of references to "devices and software operations", "information and data literacy" and "communication and collaboration". In the competency areas of "creating digital content" and "security", biology textbooks had the highest number of coding. References related to "problem solving" were found only in chemistry textbooks. Biology, physics and chemistry textbooks were not coded for "career-related competencies". The frequency distribution of total references in science textbooks according to competency areas is given in Figure 2.

Figure 2. Percentage distribution of competence areas in secondary education science textbooks

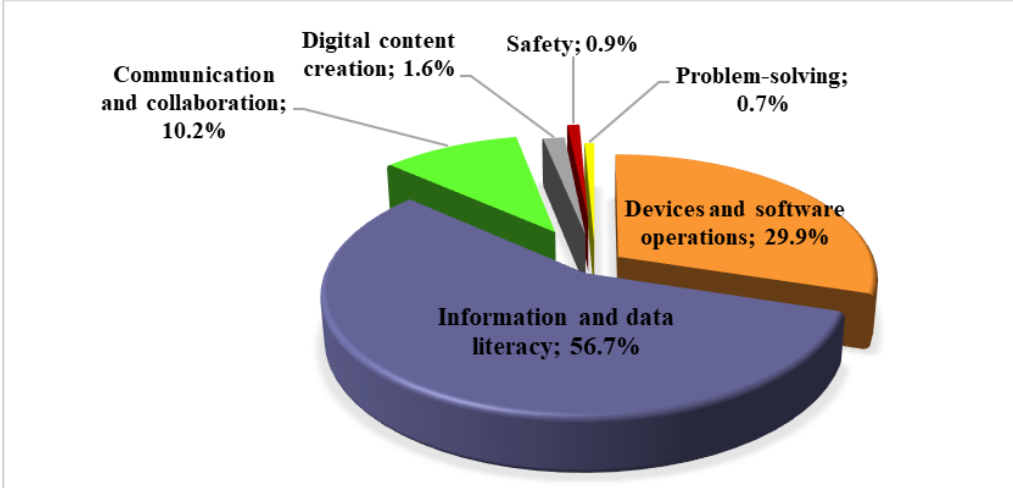


Figure 2 shows that the competences belonging to the competence area of "information and data literacy" is the area with the highest rate of coding in science textbooks with 56.7%. It is followed by "devices and software operations" with 29.9%. In secondary science textbooks, references to the competence area of "communication and collaboration" constitute 10.2% of all references. The references related to "safety" competencies in the analyzed textbooks are quite low with 0.9%. The "problem-solving" competence area is the least coded area in science textbooks with a rate of 0.7%. In none of the secondary school science textbooks, coding made within the scope of "career-related competencies".

Within the scope of the research, the presence of digital literacy competencies in textbooks according to grade levels is presented in Table 3 below.

Table 3. Competence areas and their representations in Turkish science textbooks

Competence area	9th grade		10th grade		11th grade		12th grade		Total	
	N	%	N	%	N	%	N	%	N	%
Devices and software operations	40	29.2	29	20.7	38	28.8	59	39.9	166	29.8
Information and data literacy	79	57.7	98	70.0	82	62.1	57	38.5	316	56.7
Communication and collaboration	16	11.7	6	4.3	9	6.8	26	17.5	57	10.3
Digital content creation	2	1.4	2	1.4			5	3.4	9	1.6
Safety			5	3.6					5	0.9
Problem-solving					3	2,3	1	0.7	4	0.7
Career-related competences										
Total	137	100	140	100	132	100	148	100	557	100

As can be seen in Table 3, the references to digital competence areas in textbooks increase with the increase in grade level. In the books examined, the area of competence that takes place at the highest rate at all grade levels is "information and data literacy"(56.7%).

References to the "security" and "problem-solving" competencies in the reviewed books are quite low. It can be expected that references to "problem-solving" competencies would be higher in high school textbooks because high school students are developmentally suitable for tasks demanding higher-order thinking and cognitive abilities. Additionally, it is very important to provide the new generation of students, who are digital natives, with security-related competencies such as understanding the risks and threats in the digital environment and protecting personal data and privacy. Therefore, safety competencies should be included more in science textbooks along with other textbooks. There is no content in the textbooks within the scope of "career-related competencies". These competencies are the operation of specific digital technologies, and the interpretation, and processing of data, information, and digital content. One of the overarching aims of the science curriculum is to develop career awareness and entrepreneurship skills related to science (MoNE, 2018). For this purpose, content references to career competency should be included in science textbooks according to the cognitive and affective developmental levels of the students.

Acknowledgment

This study was supported by a grant SYL-2023-10902 from the Scientific Research Projects Office of Marmara University.

References

Antoninis, M., & Montoya, S. (2018). A global framework to measure digital literacy. *Blog UNESCO*. Retrieved May 29, 2023 from http://uis.unesco.org/en/blog/global-framework-measure-digital-literacy

Erstad, O. (2013). *Digital learning lives: Trajectories, literacies, and schooling*. Peter Lang.

Erstad, O., & Voogt, J. (2018). The twenty-first century curriculum: Issues and challenges. G. A. K. Joke M. Voogt, Kwok-Wing Lai, Rhonda Christensen (Ed.), *Second handbook of information technology in primary and secondary education* (1. Baskı, s. 19-36) içinde. Springer International Publishing. https://doi.org/10.1007/978-3-319-71054-9_1

European Commission. (2022). *Guidelines for teachers and educators on tackling disinformation and promoting digital literacy through education and training* (Y. Directorate-General for Education, Sport & Culture, Eds.). Publications Office of the European Union. https://doi.org/doi/10.2766/28248

Feerrar, J., Hammer, K., & Dean, K. G., Corinne and Digital Literacy Framework Working Group. (2019). Development of a framework for digital literacy. *Reference Services Review*, *47*(2), 91-105. http://hdl.handle.net/10919/89673

Knight, B. A. (2015). Teachers' use of textbooks in the digital age. *Cogent Education*, *2*(1), 1-10.

Knight, B. A., & Horsley, M. (2011). *E-learning in higher education: Change or complementarity in the use of textbooks and other learning resources.* The Eleventh International Conference on Textbooks and Educational Media, Kaunas, Lithuani

Neumann, M. M., Finger, G., & Neumann, D. L. (2017). A conceptual framework for emergent digital literacy. *Early Childhood Education Journal*, *45*(4), 471-479.

OECD. (2005). *The definition and selection of key competencies: Executive summary*. OECD Paris.

Osterman, M. D. (2013). Digital literacy: Definition, theoretical framework, and competencies. 135-141

Payton, S., & Hague, C. (2010). *Digital literacy in practice: case studies of primary and secondary classrooms*. Futurelab

Voogt, J., & Roblin, N. P. (2010). *21st century skills*. Universiteit Twente.

Virtual Reality: Where Art meets Science

María Napal[1], Isabel Zudaire[1,], Irantzu Uriz[1], Ainhoa Méndez[1], Lander Calvelhe[2] and Javier Osés[3,4]

[1]Dpt. of Sciences. Public University of Navarre (UPNA), Spain

[2]Dpt. of Human and Education Sciences. Public University of Navarre (UPNA), Spain

[3]Dpt. Of Engineering. Public University of Navarre (UPNA), Spain

[4] Asociación de la Industria Navarra (AIN), Spain

Emerging technologies such as Augmented Reality (AR) and 3D animation enable the creation of surprising images and effects that capture the attention of the spectator, and foster comprehension of certain phenomena; for example, overcoming the limits imposed by scale to show complex and abstract concepts. This contribution describes a project (Visualtech 4.0), oriented towards linking science and technology with a didactic perspective on science and arts. High impact, artistic pieces of work are integrated into science teaching activities to be used in formal and informal teaching. To this end, optic microscopy images of materials subjected to various treatments will be merged into AR and 3D animations; this will allow unfolding the complex physical concepts involved, and adapting the didactic content to various user profiles. The project includes 4 activities, which involves a university, a research centre and various firms. It is expected to generate knowledge on STEM areas (Science, Technology, Engineering, Arts & Maths) and to foster scientific vocations, independent on the gender of the learner.

Keywords: Augmented reality, teaching innovation, materials engineering, STEAM

Introduction

The audiovisual industry has developed exponentially in the last three decades, over and over again breaking barriers in the way of displaying content, and producing audiovisual products that are increasingly spectacular and impressive for the viewer. Specially during the past decade, the film and AV sectors have gone through rapid changes due to technological and digital revolutions. Such changes have reshaped how AV content is produced, distributed and consumed (European Commission, 2024).

This opens an interesting range of possibilities for a new generation of innovative

learning scenarios based on technology, where visual representations integrated in virtual reality are combined with physical teaching resources to favor the understanding of complex, abstract and dynamic phenomena (Olympiu et al. 2013). For example, the relationship between the micro and macro scale, or the explicit or implicit interactions of physical and chemical phenomena (Eilam y Gilbert, 2014).

Benefits of the use of digital media in learning are multiple, and include the possibility (1) to generate multimodal representations (visual, auditory, tactile, textual...) of complex concepts, easing the transition between the real phenomena and the theoretical models; (2) to symbolize dynamic processes in multiple temporal and space scales; (3) to gain control over the concepts and phenomena studied (both students and teachers), which enables the emergence of timely discussion and inquiry (Tolentino et al., 2009); (4) to improve the student's interest, with an adequate stimulus to arouse curiosity in the technical subject and go deeper into the topic.

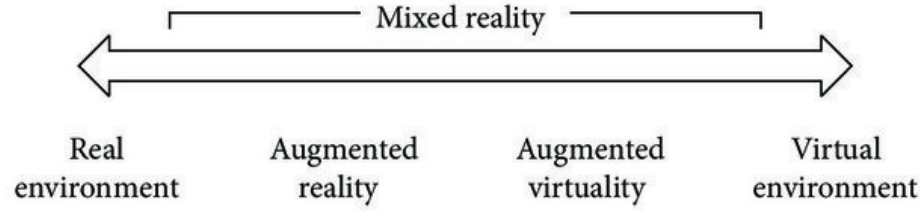
Interaction with the scale is especially interesting. Physical, chemical and biological phenomena, and the theories and laws that govern these phenomena, are different at different scales (temporal and in size). Due to that, students often present difficulties to stablish the relationship between these different levels (Becker *et al.,* 2015), although it is well proven that students must go beyond rote mathematical problem solving in order to connect their understanding of mathematical and graphical representations to the macroscopic and submicroscopic phenomena they represent. Indeed, *scale, proportion and quantity* is one of the science crosscutting concepts (NGSS, 2014), so any resource focused on it could have a multiplier and profound effect.

Among the emerging technologies that are currently setting the trend in education, probably the one with the greatest potential is Augmented Reality (AR). Augmented Reality

In fact, mixed reality technologies (Augmented Reality, Virtual Reality, Mixed Reality and other haptic alternatives) are identified as one of the influential emerging technologies, practices and trends according to the prestigious Horizon report (Pelletier *et al*., 2023).

These technologies may adopt different formats, that range from Augmented Reality (AR), which combine 3D animation with reality to 100% virtual environments, where we can move thanks to the use of immersive or not Virtual Reality tools (*Virtual Reality*, VR), or other modalities of *Mixed Reality* (MR) in the *continuum* reality – virtuality (Kahn, Johnston y Ophoff, 2019).

Figure 1: Milgram's Continuum of reality (Kahn, Johnston & Ophoff, 2019)



Pellas et al. (2021) conducted a systematic review of the decade 2010 – 2020, and found that the majority of studies in K-12 settings included in the report (n=46) reported that students were able to achieve deep learning of complex knowledge (Shi *et al.* 2019), improved their cognitive skills pertaining creativity, problem-solving and critical thinking (Chien *et al.* 2019, Segura et al. 2019, Wu *et al.* 2019), and developed better metacognition (Chang *et al.* 2018). In general, users seemed to have positive attitudes and perceptions using Virtual Reality applications even if technology limitations and potential complexity are factors that need to be addressed in the future.

On the other hand, the union between science, technology and art has been shown as a way to bring scientific knowledge closer to the general public, or to increase the interest of students from different educational levels in any of the disciplines encompassed within the acronym STEAM (Science, Technology, Engineering, Arts and Maths) (Kelly et al., 2001, Papacosta and Hanson, 1998). Many branches within contemporary art, such as digital art, virtual art or logarithmic art, explore the limits between science and art to transmit sensations and, at the same time, capture part of the technical knowledge enacted to generate the work. Various branches within contemporary art, such as digital art (Levis, 2001), virtual art (Popper, 2007) or logarithmic art (Bovill, 1996), mix the boundaries between science and art to transmit sensations and at the same time capture part of the technical knowledge used to make the work.

Despite the fact that art and science are apparently necessarily different from the point of view of the study and dissemination of knowledge, their conjunction and complementarity provide great advantages from the didactic point of view and add up in terms of their ability to impact the viewer. Science tries to understand the world based on knowledge accumulated over time, based on reproducible results and demonstrable and contrasted facts, always seeking objectivity. On the other hand, the humanities, which are the academic disciplines that look at the human condition, use primarily analytical, critical, and speculative methods.

On the one hand, art does not seem to fit in with this conception of the sciences, as it has an important component that depends on the personal experiences of the artists, their intuition or metaphors, and is often subjective and sensory. But, on the other hand, in art there are also processes of research, experimentation, generation and transmission of technical and specific knowledge that allow us to find a multitude of ways of union with science. This union between these apparently discordant worlds allows us to obtain results that surprise the viewer and manage to produce in him a response that is often not possible to produce through other approaches.

As such, and despite the fact that apparently art and science differ with regards to their methods and dissemination of knowledge, their conjunction and complementarity could be a tool to be used to bring the message in a more direct way to the receptor, even going so far as to generate STEAM vocations (Plonczak y Goetz, 2015).

Objective

In this context, our main goal was to create multimedia didactic resources based on AR and 3D animation to be implemented in workshops and exhibitions in formal and non-formal learning spaces to enable comprehension of complex scientific concepts, considering specific difficulties pertaining to the different levels of representation (submicroscopic, macroscopic and symbolic).

Future development will include measuring the impact on conceptual understanding and attitudes towards knowledge following the implementation of these didactic with specific attention to the gender dimension.

Methods

The project proceeds in four main phases.

1. Treatment to materials

   Thermic and chemical treatments were applied on materials and surfaces (aesthetic coatings, corrosion, rusting, deformation, and plasma treatments), following standard procedures to investigate the microstructure of various metallic materials and to analyze the effects of corrosion and rusting processes at high temperature.

2. Stablishing learning progressions

   The second step was to identify and develop didactic contents based on the materials, images and video obtained in the previous work package, following extensive reviews of educational literature and resources, as described in the section “results”.

3. Creating digital multimedia resources that combine 3D animations and

AR 3D models (depending on the property involved and the specific difficulty being targeted), that focused in the themes from the topics above that would most benefit from MR approaches.

4. Beyond the scope of this communication, the next step will include running workshops and exhibitions, and evaluate the impact of the resources created on different users, via questionnaires addressed to evaluate the impact on (a) comprehension; (b) interest for the study of scientific phenomena; (c) the impact of AR-enhanced resources.

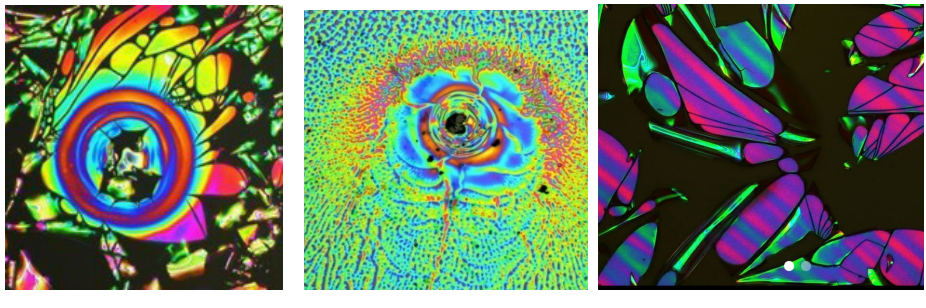
Results

In the following lines, results are presented for each of the phases.

Treatment to materials

The thermal and chemical analysis applied to several alloys allowed the research team to obtain artistic images that could later be used with aesthetic purposes (Figure 2).

Figure 2. (a) Rainbow flame. 200X micrograph Fragment of a nanometer-thick inorganic layer placed on glass and heat-treated; (b) Bloom; (c) The beauty of defects



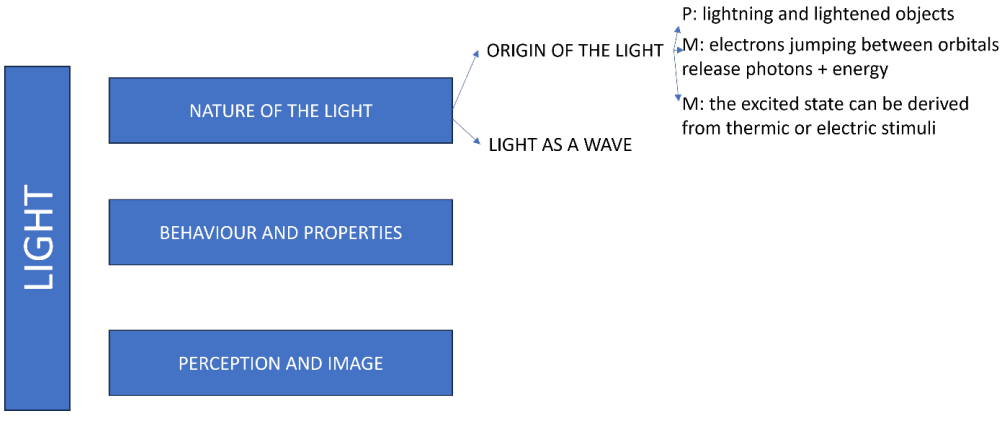
Learning progressions

The second step was to identify and develop didactic contents based on the materials, images and video obtained in the previous work package. It was estimated that these first set of generated materials could be most suitable for teaching contents related to light.

Following this realization, the work was developed as follows:

1. Map the main themes and sub-themes about light, by educational stage (Elementary, Middle, High school), in reference to the Spanish learning standards, published text books and other sources such as the NSTA's DCI progressions (NGSS Leads States, 2013) (Figure 3).

Figure 3. Example of learning progression stablished for one of the subthemes inside the first main theme. The process progress similarly for the rest of indicators.



2. Analyze the learning difficulties and learning progressions of optic phenomena in primary, secondary and high school levels (https://spark.iop.org/) and other sources.

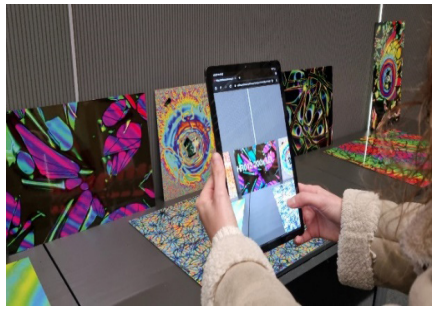
3. List of conflicts or challenges that would most benefit from an approach based on AR, via discussion with relevant experts.

Production of digital resources that integrate digital technologies, science and art.

This phase of development required the following steps. Learned lessons for each step are reported.

1. Analysis of the performance of the RA software in detecting complex images depending on several factors (distance, angle, illumination, etc.) (Figure 4).

Figure 4. Image recognition using AR



Recognition was very variable, and highly dependent on environmental conditions (light, brightness, ambient light – intensity and temperature), which call for caution when planning interventions based on AR that may act in unstructured environments or depend on users' own devices.

2. Development of the structure of the database to interact with the AR system.
3. Creation of educational multimedia resources and/or 3D models, that will be shown as started from the selected target.

   The opportunity to create specific resources stemming from specific learning objectives created materials well tailored (Figure 5), which allowed the researchers a fine-grained instruction. While it's potentially very effective, the cost involved in creating the materials in this way can be unsustainable in the long run.

Figure 5. Sequence of the animation explaining how do insects perceive shapes and colours, as compared to vertebrates or humans.



Moreover, the team realized how difficult is to avoid purely transmissive approaches, that pass on a distorted vision of science and its nature. Since the technology that is intended to be used in this project is very attractive, it is essential to overcome the "wow" effect, and to ensure that audiovisual materials are focused on the needs and objectives of the learner.

Discussion

This joint work with technological centers and companies specialized on AR and 3D animation will allow creating high impact educational resources, that

enable for an enriched, customizable experience.

Previous research has shown that AR and VR approaches are effective in closing the bridge between different levels of representation, and they may provoke deeper understanding of certain phenomena (Pellas *et al.,* 2021). Coherently, in this project the ability to enhance reality with symbolization of the abstract entities proved effective for creating resources targeting specific learning difficulties. Their effectiveness remains yet to be empirically tested.

While the integration of all the disciplines (science, arts and technology) poses a challenge on the team, whose members specialize in one of the areas, the TPACK model (Technological Pedagogical Content Knowledge) (Mishra and Koehler, 2009) suggest that the impact will depend on this integration being effective. This aspect that pertains to the conceptualization of STEM education opens also a promising follow-up line.

Last, the process of co-creating the educational resources of the project meant an opportunity to engage in high – level didactic discussions, with great potential for teachers professional development, that should be further explored.

Acknowledgements

This work has been funded by Dpto. de Universidad, Innovación y Transformación Digital del Gobierno de Navarra (PC 173-174 VISUAL TECH 4.0).

References

Becker, Nicole & Stanford, Courtney & Towns, Marcy & Cole, Renee. (2015). Translating across macroscopic, submicroscopic, and symbolic levels: The role of instructor facilitation in an inquiry-oriented physical chemistry class. Chem. Educ. Res. Pract.. 16. 10.1039/C5RP00064E.

Chang S-C, Hsu T-C, Kuo W-C, Jong MS-Y (2019b) Effects of applying a VR-based two-tier test strategy to promote elementary students' learning performance in a geology class. Br J Edu Technol. https://doi.org/10.1111/bjet.12790

Chien S-Y, Hwang G-J, Siu-Yung Jong M (2019) Effects of peer assessment within the context of spherical video-based virtual reality on EFL students' English-speaking performance and learning perceptions. Comput Educ. https://doi.org/10.1016/j.compedu.2019.103751

Eilam, B., Gilbert, J.K. (2014). The significance of visual representations in the teaching of science. In: Science teachers' use of visual representations (pp. 3-28). Springer, Cham. https://doi.org/10.1007/978-3-319-06526-7_1.

European Commision, 2024. Culture and creativity. https://culture.ec.europa.eu/cultural-and-creative-sectors/audiovisual

Khan, T., Johnston, K., & Ophoff, J. (2019). The impact of an augmented reality application on learning motivation of students. *Advances in human-computer interaction*, *2019*.

Koehler, M., & Mishra, P. (2009). What is technological pedagogical content knowledge (TPACK)?. *Contemporary issues in technology and teacher education*, *9*(1), 60-70.

NGSS Lead States. 2013. *Next Generation Science Standards: For States, By State*s. Washington, DC: The National Academies Press.

Olympiou, G., Zacharias, Z., & Dejong, T. (2013). Making the invisible visible: Enhancing students' conceptual understanding by introducing representations of abstract objects in a simulation. *Instructional science*, *41*, 575-596.

Papacosta, P., and Hanson A. (1998). Artistic expressions in science and mathematics. Journal of College Science Teaching 27: 250-252.

Pellas, N., Mystakidis, S. & Kazanidis, I. Immersive Virtual Reality in K-12 and Higher Education: A systematic review of the last decade scientific literature. *Virtual Reality* **25**, 835–861 (2021). https://doi.org/10.1007/s10055-020-00489-9

Pelletier, K., Robert, J., Muscanell, N., McCormack, M., Reeves, J., Arbino, N. and Grajek, S. (2023) EDUCAUSE Horizon Report, Teaching and Learning Edition. EDUCAUSE.

Plonczak, I. y Goetz, S. (2015), Understanding the Art in Science and the Science in Art through Crosscutting Concepts. Science Scope: 57-63

Segura R et al (2019) VR-OCKS: a virtual reality game for learning the basic concepts of programming. Comput Appl Eng Educ. https://doi.org/10.1002/cae.22172

Shi A, Wang Y, Ding N (2019) The effect of game-based immersive virtual reality learning environment on learning outcomes: designing an intrinsic integrated educational game for pre-class learning. Interact Learn Environ. https://doi.org/10.1080/10494820.2019.1681467

Tolentino, L., Birchfield, D., Megowan-Romanowicz, C., Johnson-Glenberg, M., Kelliher, A. y Martinez, C. (2009). Teaching and learning in the Mixed-Reality science classroom. Journal of Science Education and Technology 18 (6): 501 – 517. https:// doi.org/10.1007/s10956-009-9166-2

Wu J, Guo R, Wang Z, Zeng R (2019) Integrating spherical video-based virtual reality into elementary school students' scientific inquiry instruction: effects on their problem-solving performance. Interact Learn Environ. https://doi.org/10.1080/10494820.2019.1587469

(Not) Learning Alone with Augmented Reality: Continuing Results From a Comparative Study in an Undergraduate Physics Laboratory Course

Thomas Sean Weatherby[1], Sebastian Kapp[2], Michael Thees[2], Fabian Beil[2], Jan-Philipp Burde[3], Thomas Wilhelm[1] and Jochen Kuhn[4]

[1]Department for Physics Education Research,
Goethe University, Frankfurt am Main, Germany.

[2]Physics Education Research Group,
TU Kaiserslautern, Kaiserslautern, Germany.

[3]Physics Education Research Group,
Eberhard Karls University, Tübingen, Germany.

[4]Chair of Physics Education, Faculty of Physics,
Ludwig-Maximilians-Universität München, Germany.

The COVID-19 pandemic meant we spent a lot of time in video calls and remote teaching and learning. Other methods of teaching and learning, such as laboratory practical work, were not so easy to digitise from their in-person forms. Although originally conceived as an investigation into differing types of media and representational forms, infection control methods necessitated changing social forms - from pair to individual work. Carried out in summer semester 2020, this investigation looks to compare the cognitive load, learning gains and system usability of two different apparatus for the display of measurement information for a set of eight small inquiry practicals on simple DC Circuits in an undergraduate practical course for learners with physics as a minor. In the first condition ($N_{Tablet\ 2D}$=28) voltage and current information is displayed on a spatially separated dashboard style array of analogue dials for each circuit component (Tablet 2D condition). This same visualisation is used in the second condition ($N_{Tablet\ AR}$=23), however, these same virtual dials are shown in Augmented Reality (AR), floating above each real component when viewed through the tablet camera (Tablet AR condition). We report no significant differences in cognitive load perceived under either condition and no learning gains under either condition. This is the case despite "excellent" system usabilities of $M_{Tablet\ 2D}$= 82.0 (SD= 17.7) and $M_{Tablet\ AR}$= 84.6 (SD= 13.3) respectively. These findings are discussed with reference to prior work.

Keywords: Augmented Reality, physics laboratory experiments, cognitive load

Introduction

Inquiry Learning, Cognitive Load, Spatial and Temporal Contiguity

In our investigation, learners carry out a practical examining the physical laws that govern simple, direct-current circuits. The end goal of such a practical is to improve the learners' understanding of these physical laws, but the task at hand is measurement and recording of current and voltage, the interpretation of these results, as well as, physically manipulating the circuit components that make up these practical set-ups. In this mode of inquiry learning meaning-making and understanding is emergent from the tasks and social interactions surrounding them (de Jong, 2019) and the combination of virtual and real experiments is essential (Wörner et al., 2022). As these cognitive gains are made on this secondary level, it is important that cognitive resources are made available in order to facilitate this. We design the two conditions considering multimedia learning theory (Mayer & Fiorella, 2014) avoiding split-attention and considering contiguity principles (displaying information at relevant times, in relevant places); we measure the usage of cognitive resources using cognitive load theory (Sweller, 1988; van Merriënboer & Sweller, 2005). Cognitive Load theory divides the types of cognitive load into three: extrinsic (ECL), intrinsic (ICL) and germane (GCL). ECL is the (reducible) cognitive load that is due to the instructional materials. ICL is the cognitive load intrinsic to the task. GCL is the cognitive load caused by the creation of new schemata, the kind of load to be maximised for learning. We measure this with an adapted version of the CLS scale by Leppink et al. (2013, 2014) by Thees et al. (2020).

Learning Context and Investigation

We present here a randomised controlled trial in a graded physics practical laboratory course for (N = 51) biology undergraduates at a German university. The topic under instruction is simple DC circuits and the laws that govern them. The aim of the practical is that, through measuring the currents and voltages in several circuits, learners improve their conceptual understanding. Learners are assigned to one of two experimental groups, both groups have circuit components, which provide the values of the current through them, and the voltage over them, digitally to a tablet. Both groups (Tablet 2D and AR) have all the measurement information displayed to them using rotary dials on a single screen, facilitating the principal of temporal contiguity. The software for the first group (Tablet 2D) simply displays an array of virtual rotary dials in a dashboard layout. Whereas, the second group (Tablet AR) use the tablet like a viewfinder to see the measurement information (visible as identical, virtual rotary dials) placed above the responding real component, facilitating spatial contiguity of measurement information and object.

Table 2. Demographic data of each group. Gender is voluntarily disclosed.

	Tablet 2D (n = 28)	Tablet AR (n = 23)
Gender		
Female	20	14
Male	8	8
No Answer	-	1
Age		
M	20.6	19.7
SD	2.7	1.3

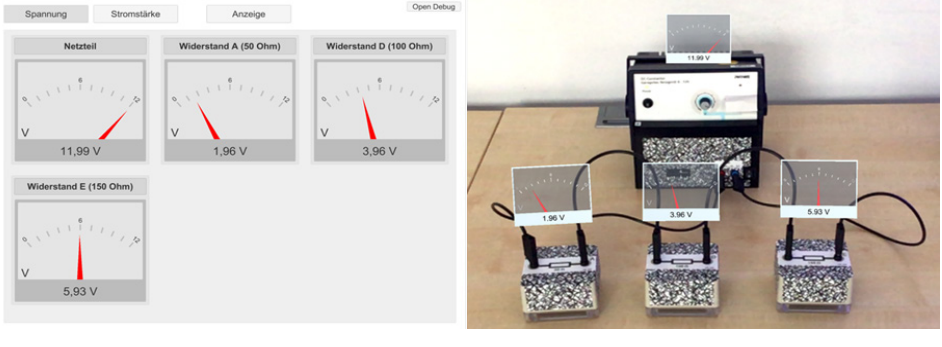
In both groups, learners work individually, due to COVID-19 transmission reduction measures, despite the practical usually being carried out in pairs. We test for conceptual knowledge gain, cognitive load, system usability as well as assessing the learners' emotional state, motivation and some demographic data.

Procedure and Materials

Software and Hardware

Measurement data (current and potential difference), from specially designed components (Kapp et al., 2022) sent via Bluetooth to the iPad, was visualised in one of two ways using the Unity 3D engine. The first of which (Tablet 2D) is a dashboard showing the measurements for all components, shown on the left of Figure 1. These same style dials are then used to visualise the information in the second application (Tablet AR). The difference here being they are shown as projections in 3D-space above the components, shown on the right of Figure 1.

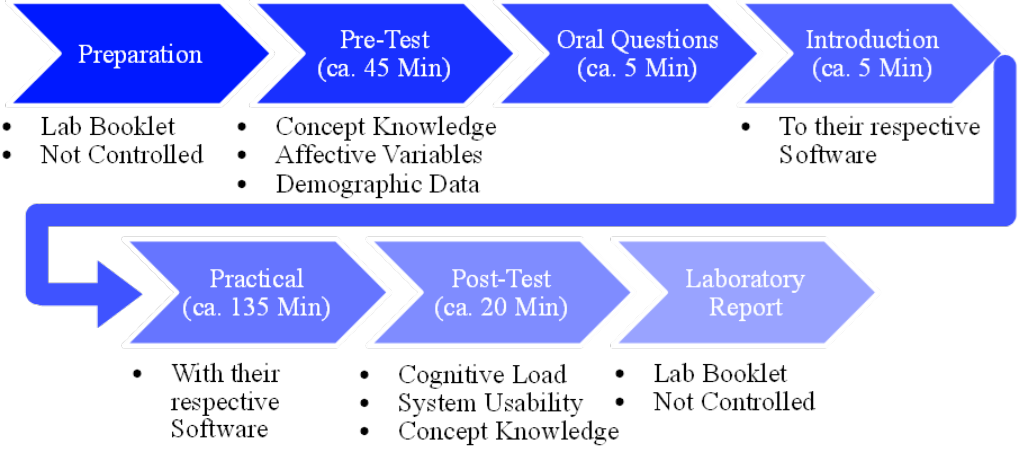
Figure 1. Software used in conditions "Tablet 2D" (left) and "Tablet AR" (right) respectively, compare Kapp et al. (2020).



Procedure

Learners have all attended a lecture series teaching the topic of DC circuit theory. Before the lab, they are provided with a booklet with which to prepare for the practical course, which they do at home. On entering the teaching laboratory, learners complete a pre-test measuring their initial concept knowledge, some demographic data and a number of affective variables, using the experience sampling method (Schallberger, 2005). The supervisors then assess the learners' preparedness to do the practical orally and give them any safety information. This is then followed by an introduction to each group's respective software. The learners then carry out their practical work, followed by a post-test, looking again at their concept knowledge, but now also measuring subjective data regarding system usability and cognitive load from their interaction with the learning media. Learners then write a laboratory report on which they are assessed.

Figure 2. A flowchart showing the order and approximate timings during the study.



Instructions and Task

Learners begin by building a circuit with two identical resistors, followed by three identical resistors and then three different resistors, all in parallel. The same sequence is then used for series circuits. The reasoning for having parallel followed by series is to centre potential difference as the principal variable of interest (Burde et al., 2022). Learners then construct two mixed circuits with switches. One in which the switch must be closed to engage a parallel circuit and a second in which a switch is used to short a resistor. For each of these circuits, learners are asked to vary the voltage from the voltage supply and record qualitative statements about the potential difference across and the current through each component, as well as making quantitative measurements.

Subjective Cognitive Load Scale

We used the same modified German version of the cognitive load scale (CLS) by Leppink et al. (2013, 2014) adapted by Thees et al. (2020) to measure intrinsic, extrinsic and germane cognitive load in 12 items on a six-point Likert scale.

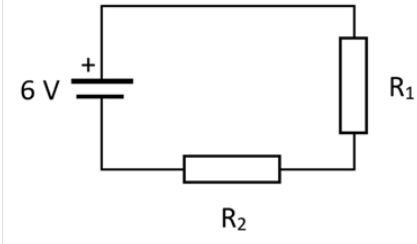
System Usability Scale

The usability of each set of measurement equipment (software and hardware) is assessed using the System Usability Scale (SUS) by Brooke (1996). Learners rate their agreement with ten statements on a five-point Likert scale (assigned values zero to four), these numbers are then summed and multiplied by two point five, hence achieving a mark out of a maximum one-hundred. The statements were translated into German from the original English and the term "system" was made more definite as "the interaction between the digital assistive system, the corresponding software application, and the experimental equipment".

Concept Knowledge Test

In order to measure if learning had taken place during the activity, a concept test is used, based upon Ivanjek et al. (2021). There are a total of 15 items on the topics of current, voltage and resistance and their interrelation in simple direct-current electric circuits. Each item consists of two parts: a 'what' part, in which the learner must make a single choice, describing a given physical quantity in the circuit shown in the corresponding diagram. The second part is a 'why' part, in which the learner also makes a single choice, justifying their reasoning. Only if both parts are answered correctly would the concept be thought of as understood in the context shown. Hence, both parts must match the physical idea for the question to be marked as correct. The adaptations taken were, firstly, to replace lightbulbs in items to resistors to be more applicable to our learning context. This necessitated not asking about their brightness, but directly about current. This may be advantageous, as we do not wish to test the learning objective that brightness and current are related in a lightbulb, as it is not obvious to learners (Maichle, 1980). Secondly, we also added some answer options to cover a fuller range of possible misconceptions and to keep the answers from being apparent to our possibly more 'testwise' undergraduates. This can be done without needing to worry about the additional reading, as the test is originally developed to be suitable for lower secondary school learners, but our undergraduates should be able to deal with the increased demands on their reading and working memory.

Figure 3. A sample item with two tiers, based on item 2 from Ivanjek et al (2021).

The circuit shown consists of a 6 V battery and two identical resistors.

a. Compare the current through both Resistors.
☐ The current through R_1 is larger than through R_2.
☐ There is a current through R_1 but not through R_2.
☐ The current through R_2 is larger than through R_1.
☐ There is a current through R_2 but not through R_1.
☐ The current through both resistors is the same.
b. Why? Choose the answer that fits your explanation best.
☐ The current is the same everywhere in the circuit.
☐ R_2 uses up some of the current, so there is less left for R_1.
☐ R_2 uses up all of the current, so there is none left for R_1.
☐ The current is shared equally between R_1 and R_2.
☐ R_1 uses up some of the current, so there is less left for R_2.
☐ R_1 uses up all of the current, so there is none left for R_2 .

Previous Work

We build upon our previous work by attempting to overcome the HoloLens' limitation of its small field of view (FOV), using instead an iPad, able to show AR overlays on the entire camera FOV. The HoloLens' restrictive FOV made it difficult to see all the measurement information simultaneously, when working on the experiment. Hence, violating the principle of temporal contiguity. Another possible case is that all components were in close proximity, so that all the information could be seen at once, however this overcrowding of information can result in spatial contiguity failure (Beege et al., 2019), again resulting in increased ECL.

In this study both experimental groups use the same display medium (an iPad), the same visualisations and complete the same tasks. Essentially here we have a replication study of Altmeyer et al. (2020) with a similar cohort, non-specialist undergraduates, as well as two key changes and one minor one. The first key change is the use of Ivanjek et al. (2021) to test concept knowledge instead of Altmeyer's use of Urban-Woldron & Hopf (2012) as modified by Burde (2018). Secondly, the difference in social form, from pair work to individual work, due

to the COVID-19 pandemic. The final is that a small change has been made to the instructions given to the learners, where Altmeyer asked single choice questions, learners here must write down their observations. In additions to the tasks Altmeyer's learners completed (series and parallel circuits) our learners must build and analyse two more complex circuits that involve switches. These differences, however, do not impede us from using the same arguments and investigating the same hypotheses as Altmeyer and our previous work, as outlined in a (since withdrawn, due to social form change) pre-registration (https://osf.io/5pw2x). Synthesising this, we form our hypotheses.

Hypotheses

In the light of less crowding of information, preventing spatial contiguity failure (Beege et al., 2019), in addition to being able to maintain both spatial and temporal contiguity of measurement and components. We formulate the first hypothesis:

Hypothesis 1: Working with the Tablet based AR measurement will result in a lower ECL than working with the Tablet Dashboard.

Following the logic that a reduced ECL will enable an increased GCL, we expect an increase in learning outcomes, as measured through concept knowledge, reflected in the second hypothesis.

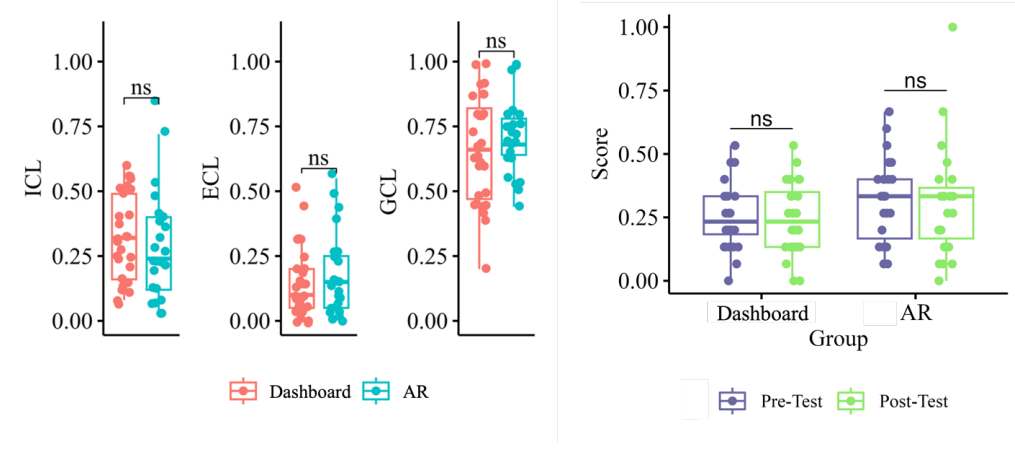
Hypothesis 2: Working with the Tablet based AR measurement software will result in a higher learning gain than working with the Tablet Dashboard measurement software.

Table 2. Summary of hypotheses and whether they are supported by previous work. Asterisks show where the caveat of spatial contiguity failure must be considered.

Hypothesis (in short)	Dependant Variable	Supported by…	
		Altmeyer et al. (2020)	Kapp et al. (2020)
H1) Less ECL for "Tablet AR" supported lab work	Subjective ECL rating	No	No*
H2) Higher learning gains for "Tablet AR" supported lab work	Concept Knowledge	Yes	No*

Results

Figure 2. Left: Boxplots showing comparisons of intrinsic, extrinsic and germane cognitive loads in the two experimental conditions. Right: Boxplots showing conceptual knowledge score comparisons pre and post under the two conditions.



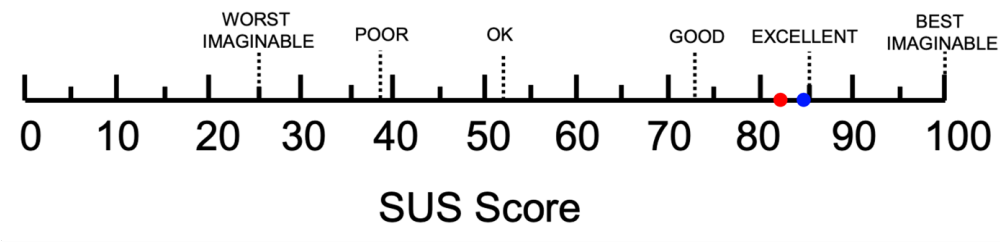
Cognitive Load

No significant differences are reported in any of the three types of cognitive load, as shown on the left of Figure 2. Significance thresholds were unmet in unpaired t-tests for comparisons of intrinsic (p=0.59), extrinsic (p=0.51) and germane (p=0.38) cognitive loads in the two experimental conditions, with two good internal consistencies and one questionable, α_{ICL}=0.81, α_{ECL}=0.63 and α_{GCL}=0.82. Hence, we find no support for the hypothesis, **H1**.

System Usability Score (SUS)

Measured using the scale from Brooke (1996), System Usability for both experimental conditions would be rated as "excellent" in accordance with the review from Bangor (2009). The non-AR and AR conditions scoring 82.0 and 84.6 respectively, with a good internal consistency of α_{SUS}=0.89. These are judged to not be sufficiently different in an unpaired t-test (p=0.55).

Figure 4. System Usabilities of the "Tablet 2D" and the "Tablet AR" conditions shown in red and blue respectively on the adjective scale from Bangor (2009).



Conceptual Knowledge

The conceptual knowledge test used to assess learning gains is a two-tiered test developed by Ivanjek et al. (2021). No significant learning gains are measured under either condition, as shown on the right of Figure 2. The only significant difference being that of "Group" and not of "Time", RM-ANOVA, $F(1,98)=0.04$. Hence, we find no support for the hypothesis, H2.

Table 2. Summary of means and standard deviations (shown in parentheses) of variables measured in this study and our previous work from Kapp et al. (2021) for comparison. Concept Knowledge scores are shown here as the outcome of the two-tiered test. Red and blue highlighted values show pre- and post- Concept Knowledge scores in individual and pair work i.e. infection control and standard conditions, respectively.

		Cognitive Load			Concept Knowledge		SUS
		ICL	ECL	GCL	Pre	Post	
Summer 2020 Individual (This Study)	Tablet 2D (n = 28)	.32 (.17)	.15 (.13)	.66 (.21)	.25 (.12)	.25 (.14)	82.0 (17.7)
	Tablet AR (n = 23)	.29 (.21)	.18 (.16)	.71 (.15)	.32 (.17)	.31 (.22)	84.6 (13.3)
Winter 2019 Pair Work (Kapp et al., 2021)	Tablet 2D (n = 29)	.22 (.15)	.17 (.17)	.67 (.20)	.29 (.13)	.38 (.24)	87.2 (12.6)
	Hololens AR (n = 27)	.28 (.16)	.17 (.12)	.60 (.19)	.31 (.12)	.35 (.20)	69.9 (17.7)

Further Explorative Results and Discussion

Our results in this study show an interesting difference to those in the results of our previous. In this study, when working alone, learners achieved no learning gains. Whereas, in Kapp et al. (2021) using the same assistive system (Tablet 2D) in pair work, learners made significant learning gains, see Table 2. There are some differences in key indicators. The intrinsic cognitive load (ICL) is judged by learners to be significantly higher when working alone ($p = 0.022$). Learners also are seen to have significantly higher negative activation (Schallberger, 2005), i.e. are more stressed and nervous, in the working alone condition ($p = 0.045$). The origins of these factors are, of course, difficult to pin down as they may have their origins in either the pandemic conditions or simply from working alone, among others. Although we do not wish to ascribe any causal links, as we did not construct our experiment to examine these differences or control all the necessary variables, we would like to show this as a promising angle for further systematic study and development of a Collaborative Cognitive Load Theory (Janssen & Kirschner, 2020), possibly with an affective-emotional component.

Acknowledgements

The authors would like to acknowledge the financial support for development and acquisition of both hardware and software from the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung; BMBF) via the projects “GeAR” (Grant No. 01JD1811B) and “gLabAssist” (Grant No. 16DHL1022). We thank the division of Microelectronic Systems Design at the Technische Universität Kaiserslautern (chaired by Prof. Norbert Wehn), especially Frederik Lauer and Carl Rheinländer, for providing the hardware components that the learners used.

References

Altmeyer, K., Kapp, S., Thees, M., Malone, S., Kuhn, J., & Brünken, R. (2020). The use of augmented reality to foster conceptual knowledge acquisition in STEM laboratory courses—Theoretical background and empirical results. British Journal of Educational Technology, 51(3), 611–628. https://doi.org/10.1111/bjet.12900

Bangor, A. (2009). Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. Journal of User Experience, 4(3), 114–123.

Beege, M., Wirzberger, M., Nebel, S., Schneider, S., Schmidt, N., & Rey, G. D. (2019). Spatial Continuity Effect vs. Spatial Contiguity Failure. Revising the Effects of Spatial Proximity Between Related and Unrelated Representations. Frontiers in Education, 4. https://www.frontiersin.org/articles/10.3389/feduc.2019.00086

Brooke, J. (1996). SUS: A ‘Quick and Dirty’ Usability Scale. In Usability Evaluation In Industry. CRC Press. https://doi.org/10.1201/9781498710411

Burde, J.-P. (2018). Konzeption und Evaluation eines Unterrichtskonzepts zu einfachen Stromkreisen auf Basis des Elektronengasmodells. Studien Zum Physik- Und Chemielernen, 259. https://doi.org/10.30819/4726

Burde, J.-P., Weatherby, T. S., & Wilhelm, T. (2022). Putting Potential at the Core of Teaching Electric Circuits. The Physics Teacher, 60(5), 340–343. https://doi.org/10.1119/5.0046298

de Jong, T. (2019). Moving towards engaged learning in STEM domains; there is no simple answer, but clearly a road ahead. Journal of Computer Assisted Learning, 35(2), 153–167. https://doi.org/10.1111/jcal.12337

Ivanjek, L., Morris, L., Schubatzky, T., Hopf, M., Burde, J.-P., Haagen-Schützenhöfer, C., Dopatka, L., Spatz, V., & Wilhelm, T. (2021). Development of a two-tier instrument on simple electric circuits. Physical Review Physics Education Research, 17(2), 020123. https://doi.org/10.1103/PhysRevPhysEducRes.17.020123

Janssen, J., & Kirschner, P. A. (2020). Applying collaborative cognitive load theory to computer-supported collaborative learning: Towards a research agenda. Educational Technology Research and Development, 68(2), 783–805. https://doi.org/10.1007/s11423-019-09729-5

Kapp, S., Lauer, F., Beil, F., Rheinländer, C. C., Wehn, N., & Kuhn, J. (2022). Smart Sensors for Augmented Electrical Experiments. Sensors, 22(1), 256. https://doi.org/10.3390/s22010256

Kapp, S., Thees, M., Beil, F., Weatherby, T., Burde, J.-P., Wilhelm, T., & Kuhn, J. (2020). The Effects of Augmented Reality: A Comparative Study in an Undergraduate Physics Laboratory Course. Proceedings of the 12th International Conference on Computer Supported Education - Volume 2: CSEDU, 2020, 197–206. https://doi.org/10.5220/0009793001970206

Leppink, J., Paas, F., Van der Vleuten, C. P. M., Van Gog, T., & Van Merriënboer, J. J. G. (2013). Development of an instrument for measuring different types of cognitive load. Behavior Research Methods, 45(4), 1058–1072. https://doi.org/10.3758/s13428-013-0334-1

Leppink, J., Paas, F., van Gog, T., van der Vleuten, C. P. M., & van Merriënboer, J. J. G. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. Learning and Instruction, 30, 32–42. https://doi.org/10.1016/j.learninstruc.2013.12.001

Maichle, U. (1980). Verstehens- und Lernprozesse im Elektrizitätslehreunterricht der Sekundarstufe I aus kognitionspsychologischer Sicht. Der Physik-Unterricht, 14(4), 7–15.

Mayer, R. E., & Fiorella, L. (2014). Principles for Reducing Extraneous Processing in Multimedia Learning: Coherence, Signaling, Redundancy, Spatial Contiguity, and Temporal Contiguity Principles. In R. E. Mayer (Ed.), The Cambridge Handbook of Multimedia Learning (2nd ed., pp. 279–315). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.015

Schallberger, U. (2005). Kurzskalen zur Erfassung der Positiven Aktivierung, Negativen Aktivierung und Valenz in Experience Sampling Studien (PANAVA-KS) (6; 'Qualität des Erlebens in Arbeit und Freizeit' Untersuchungen mit der Experience Sampling Method, p. 87). Psychologisches Institut der Universität Zürich.

Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. Cognitive Science, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Thees, M., Kapp, S., Strzys, M. P., Beil, F., Lukowicz, P., & Kuhn, J. (2020). Effects of augmented reality on learning and cognitive load in university physics laboratory courses. Computers in Human Behavior, 108, 106316. https://doi.org/10.1016/j.chb.2020.106316

Urban-Woldron, H., & Hopf, M. (2012). Entwicklung eines Testinstruments zum Verständnis in der Elektrizitätslehre. Zeitschrift Für Didaktik Der Naturwissenschaften, 18, 201–227.

van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions. Educational Psychology Review, 17(2), 147–177. https://doi.org/10.1007/s10648-005-3951-0

Wörner, S., Kuhn, J., & Scheiter, K. (2022). The best of two worlds: A systematic review on combining real and virtual experiments in science education. Review of Educational Research, 92(6), 911-952. https://doi.org/10.3102/00346543221079417

Digital Teaching Scenarios on Advanced STEM Topics for Higher Education

Emily Michailidi[1], Eilish McLoughin[2], Argyris Nipyrakis[1,3], Athanasia Kokolaki[1], Gunnar Friege[4], Priit Reiska[5], & Dimitris Stavrou[1]

[1]University of Crete, Greece, [2]Dublin City University, Ireland, [3]University of Groningen, The Netherlands, [4]University of Hannover, Germany, [5]Tallinn University, Estonia

The STEM Digitalis project's aim is to develop approaches which are effective for teaching STEM in a blended learning environment. Five STEM scenarios were identified as offering the potential to engage pre-service teachers with advanced STEM topics, to facilitate the development of their conceptualisation of STEM learning and to provide them with learning experiences which support them to develop their pedagogical content knowledge around STEM. This workshop, responding to the need for developing blended and distance learning environments for teaching advanced STEM topics in higher education, aims to familiarize participants with the educational use of various digital resources for science teaching in online and distance learning contexts and provide examples of good practices regarding teaching and learning strategies that promote meaningful use of digital technologies for teaching STEM topics in blended or distance learning environments.

Keywords: Digital tools, STEM education, higher education.

Background and Rationale

The COVID-19 pandemic forced institutions worldwide to quickly transition instruction from classrooms to configurations of distance learning. The whole enterprise was new to many academic institutions and to many staff members and students as well, with the consequence that they were not properly equipped with the necessary methodology for blended or distance education. Additional difficulties arose in the teaching of science courses that require students' hands-on interaction for them to develop not only scientific knowledge but also inquiry and practical skills for the operation of scientific equipment, data collection and processing etc. The closure of universities over a long period raised questions about the future of higher education and about the pandemic (or other crisis) preparedness of universities, in terms of whether in the future they will be in position to switch to an accessible, inclusive, equitable and engaging science education online with proper support for students and staff members (Dillon &

Avraamidou, 2020).

In this context, many efforts have been made over the last years towards developing online or blended digital resources and teaching and learning environments for tertiary science education that may be used in order to extend the capabilities of science education in cases where physical presence is not feasible. From a pedagogical perspective, online teaching has a positive impact on teaching and assessment strategies (Cook & Steinert, 2013). Teachers can incorporate effective pedagogical and instructional strategies such as games, interactive models, computer simulations and animations for learners to engage in meaningful knowledge construction. The multimodality and availability of these rich resources are advantageous as online digital resources can enrich the classrooms and improve student learning, by providing high-quality and collaborative online learning experiences synchronously and asynchronously (Eichler & Peeples, 2013).

The STEM Digitalis Project

Aims and Goals

Recognizing the aforementioned contextual needs and taking into consideration the affordances of distance or blended digital resources for science education, the STEM Digitalis Erasmus+ project (https://stemdigitalis-project.eu/) aimed at developing blended and distance learning environments for teaching advanced STEM topics to prospective primary and secondary education science teachers. The ultimate goal is to digitally equip prospective science teachers not only regarding the use of digital tools (e.g. educational robotics, data loggers, virtual and augmented reality, etc.) but also regarding the effective integration of such tools in teaching and learning processes of advanced STEM topics.

STEM Digitalis' Digital Scenarios

Five STEM scenarios were identified as offering potential to engage pre-service teachers with advanced STEM topics, to facilitate the development of their conceptualisation of STEM learning and to provide them with learning experiences which support them to develop their pedagogical content knowledge around STEM. Each scenario is presented over several units, with each unit comprising of several learning activities. These activities were designed around a five-dimensional instructional framework, depicted in Figure 1, which was developed in the STEM Digitalis project, informed by literature on integrated STEM learning (Radloff & Guzey, 2016) and blended learning (Margulieux et al., 2016). The activities within each unit were designed such that educators could choose subsets of the activities to make up a learning trajectory best suited to their programme's context. Sample trajectories were visualised using

the LePlanner programme and published using the STEM Digitalis Pressbooks platform.

Figure 1. Five-dimensional instructional framework for STEM Digitalis scenarios.



The content of each scenario was trialed in local educational contexts by each of the project members. Based on feedback from these trials, the content of the five scenarios was amended and refined. Further refinement happened following implementation of the scenarios during the STEM Digitalis Summer School (which took place in Crete in June 2022) and the STEM Digitalis Online Joint Seminars (which took place in January - September 2023).

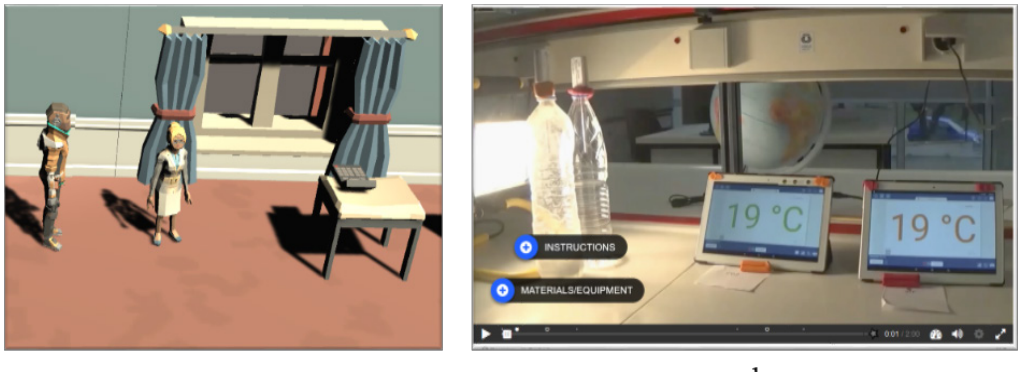
During the ESERA 2023 workshop participants implemented three of the digital teaching scenarios developed in the context of the STEM Digitalis project, engaging with different types of educational digital tools in the context of three advanced STEM topics, namely Climate change, Water, and Ocean batteries & energy farms.

Digital scenario on Climate Change

The rapid climate change that is observed during the last century has important impacts on the natural environment, threatening the biodiversity and survival of ecosystems as well as human societies, affecting health and economies. Key leverage point to achieve the goal of citizens' climate literacy is climate education that may equip citizens with knowledge on climate change causes, consequences, and mitigation-adaptation measures. The digital teaching scenario of climate change consists of 3 units that focus on (i) drivers of climate change, (ii) carbon emissions generated by anthropogenic activities and mitigation measures taken regarding renewable sources of energy and (iii) the negotiation of a

socioscientific issue regarding the lignite phase out to meet EU goals for climate neutrality. In order to cover these aspects of climate change diverse digital tools have been used such as: Unity 3D, mobile apps, H5P, online communication and collaboration tools. Particularly, the first unit on climate change drivers employs a Unity 3D-based serious game (shown in Figure 2a) through which participants collect and compare data concerning temperature and carbon dioxide rise through decades of 1980s until 2010s in order to form conclusions about global temperature rise and carbon dioxide average concentration rise. Participants also explore the correlation between the global temperature rise and the carbon dioxide concentration rise via an integrated interactive experiment with integrated hotspots that provide additional information, multiple choice questions and links to online dashboards for sharing conclusions, shown in Figure 2b.

Figure 2. Snapshot of the serious game (a) and the interactive video experiment (b) environment



a b

The second unit is based on interactive maps, depicted in Figure 3a, through which participants explore data on the amounts of CO_2 emitted per anthropogenic activity per country and on a mobile application, shown in Figure 3b, through which participants calculate their personal CO_2 emissions.

Figure 3. Snapshot of the interactive map (a) and the mobile application (b)



a b

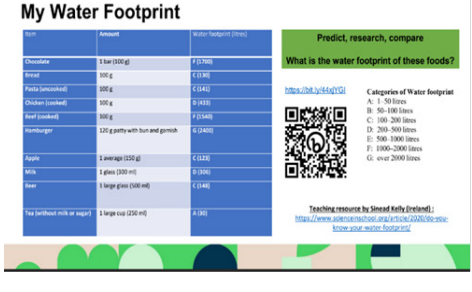
The third unit is focused on the negotiation of perspectives of multiple stakeholders on a socioscientific issue raised in the context of the need for climate neutrality. In order to do so, participants take part in a "treasure hunt" supported by an augmented reality application where they scan an image with their mobile devices and information appears on the screen. By solving riddles and using maps, participants in small groups have to trace the image of one of the stakeholders that are involved in the debate.

Digital scenario on Water

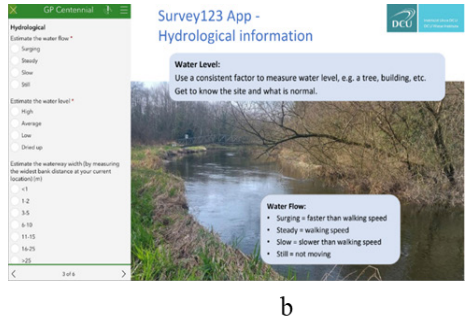
In a world faced with the demands of a growing population, increased urbanization, and climate change, the world's water resources are under threat, with Global water demand set to increase by 22% by 2030. This highlights the urgent need for research, and the finding of technological solutions to combat our growing water problems, national and internationally. Sustainable water means a nation that can be water self-sufficient: ensuring there is enough water to meet multiple needs, from agriculture to municipal and industrial. This scenario focuses on developing learners' understanding of how individuals and societies need to work together to provide safe, secure drinking water for all. These resources and activities help us to think about how we use water and how we might use a little less. The digital teaching scenario of water consists of 3 units that focus on developing understanding of (i) water conservation and the water cycle, (ii) an individual's water footprint and the role of citizen science in water education and (iii) water monitoring and data analysis. This scenario promotes the use of digital communication and collaboration tools, CODAP web-based software for data analysis, CMAP Concept mapping software, mobile apps, online forms and databases.

The first unit facilitates learners to prepare causal explanations and visual representations of the water cycle. As shown in Figure 4a, learners calculate and reflect on their own water footprint based on daily food consumption.

Figure 4. Snapshots of water footprint activity (a) and water quality activity (b)



a



b

The second unit focuses on the notion of data literacy and its importance in terms of social justice in the 21st century as well as on the development of skills in graphing and interpreting water data sets and observations. Particularly learners are presented with a large dataset, recorded via a multi-sensor sonde in Dublin port and asked to interrogate the dataset using CODAP.

The third unit introduces learners to the opportunities and benefits of being involved in co-created citizen science projects around water, as shown in Figure 5. Specifically, learners work in groups to collect water samples, to test them and to upload their findings to the Freshwater Watch project database.

Figure 5. Snapshots of the citizen science activities



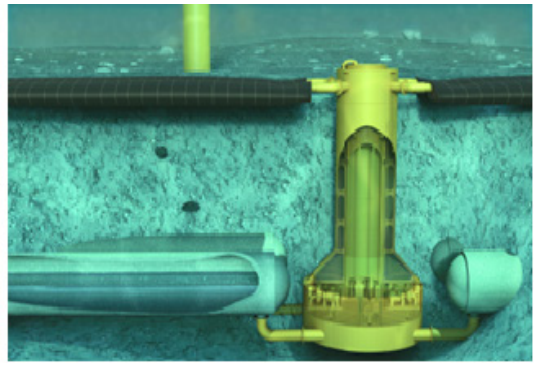
a b

Digital scenario on Ocean Batteries & Ocean Energy Farms

The growing need for sustainable and renewable energy sources appears as an eco-friendly solution to the energy crisis which is a contemporary global problem. However, fluctuations appear in the energy supply from renewable energy sources due to the instability and unpredictability of local weather conditions (e.g. wind, sunlight). Therefore, it appears an additional need for sustainable and eco-friendly energy storage techniques that would support the operation of renewable energy farms. In parallel, research on off-shore energy farms consisting of wave energy converters and wind turbines aims for higher efficiency levels in energy harvesting. In light of the above, Ocean Grazer (www.oceangrazer.com) is a hybrid renewable energy harvester, which means that it integrates both energy production and energy storage modules (DiModugno, 2021). Indicative illustrations of such a module are presented in Figure 6. Specifically, the installation combines wave energy converter technology (Asiikkis et al., 2023) with wind turbines and on-site energy storage devices called ‘ocean batteries’ in order to harvest energy off-shore (Vissers, 2019). Ocean batteries are devices that store energy in the form of potential energy by making use of differences in hydrostatic pressure between an internal rigid reservoir kept in low (atmospheric) pressure and a flexible bladder kept in high pressure at the bottom of the installation inside the ocean.

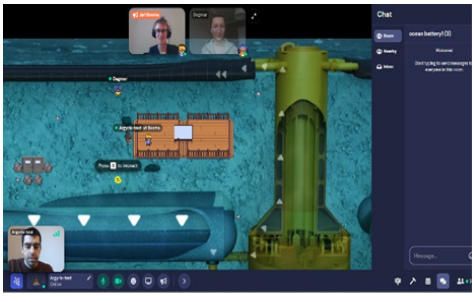
Figure 6. Illustrations from the Ocean Grazer μodule (oceangrazer.com)



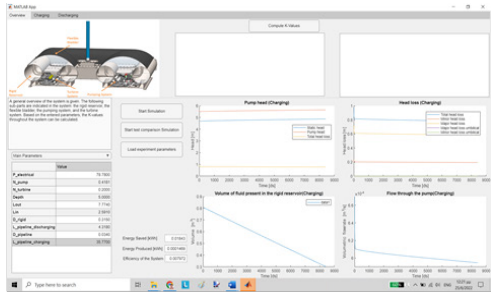


This teaching module, based on the aforementioned phenomena and applications, engages students with engineering practices on exploring, testing, revising, and reflecting on a simulation model on ocean batteries and ocean energy farms, while they collaborate with peers through an online gamification platform. In specific, the module comprises three submodules. In the first submodule, students log into a customised gamification environment (gather.town) as shown in Figure 7a, formulate small groups online, and go through a historical overview of renewable energy sources. In order to proceed among stages in the platform, the users interact with virtual assistants through Augmented Reality mobile apps (Metaverse app), who also assess their knowledge in each stage in order to unlock the passwords needed to proceed to the next stage. In the second submodule, users engage in engineering design practices by carrying out several variable testing procedures through a Matlab-based simulation programme (Ocean Grazer), as shown in Figure 7b, in order to optimise the design of i) ocean batteries (charging and discharging features), ii) wave energy converters, iii) wind turbines. In the third submodule, students reflect on their design decisions on the ocean batteries and energy farm module, as well as on interdisciplinary, epistemological, and socioscientific aspects related to the activities they previously experienced.

Figure 7. Snapshot of the gamification platform (a) and of the ocean battery simulation (b)



a



b

Key learnings from digital scenarios' implementation

Based on the implementation of the digital scenarios, engaging participants in student-centred and inquiry-based approaches emerges as crucial to integrating technology for STEM teaching and learning. The design and implementation of the scenarios provided some exemplar ways in which technology can be used as an indispensable part of the teaching scenario e.g., in order to collect and analyse data and draw inferences that can lead to achieving the teaching goals.

Moreover, technology can be used as a complementary asset of the educator by aiding in an individualised and flexible manner, e.g., by providing content- or process-related assistance and feedback to participants to accomplish the tasks. However, assistance and guidance from the teacher were deemed important in the implementation of the scenarios; hence we can conclude that a combination of assistance from both educators and technology can be the most productive method.

Additionally, the implementation of contemporary and real-world topics can provide a fruitful context that can stimulate participants' interest and engagement with the content, phenomena, and applications. In this light, taking into consideration sustainability goals such as aiming at affordable and clean energy and climate action can activate meaningful integration of content topics that relate to participants' everyday life. Also, engaging participants with socioscientific issues that are relevant to the topic can activate an additional metalevel of participants' understanding and engagement with the topic. However, discussions about the impact of technologies are recommended to be assisted by educators and experts in order to discuss more deep insights about the topic.

Finally, collaboration with peers, educators and experts is an important factor that can enhance learning and engagement with the topic, as well as it can assist in overcoming difficulties in using technology. However, adequate interaction across different groups of participants is needed to take into consideration

multiple views. In this direction, technological tools and platforms can provide an environment that can assist collaboration in distance and blended learning environments.

References

Asiikkis, A., Grigoriadis, D., & Vakis, A. (2023). Maximizing the Surge Amplitude of a Floater Through an Adaptable Mooring Tightening Technique. *Proceedings of the European Wave and Tidal Energy Conference, 15*. https://doi.org/10.36688/ewtec-2023-274

Cook, D. A., & Steinert, Y. (2013). Online learning for faculty development: A review of the literature. *Medical Teacher, 35*(11), 930–937. https://doi.org/10.3109/0142159X.2013.827328

Di Modugno, A. (2021). *A Techno-Financial Model of Ocean Grazer's Hybrid Integration Proposing the Ocean Battery as Deployable On-site Storage Profitability Optimizer for Offshore Renewable Production Plants.* [Master Design project, University of Groningen].

Dillon, J., & Avraamidou, L. (2020). Towards a viable response to COVID-19 from the science education community. *Journal for Activist Science & Technology Education, 11*(2), 1–6. https://doi.org/10.33137/jaste.v11i2.34531

Eichler, J. F., & Peeples, J. (2013). Online homework put to the test: A report on the impact of two online learning systems on student performance in general chemistry. *Journal of Chemical Education, 90*(9), 1137–1143. https://doi.org/10.1021/ed3006264

Margulieux, L.E., McCracken, W.M. & Catrambone, R., (2016). A taxonomy to define courses that mix face-to-face and online learning. *Educational Research Review, 19*, 104-118. https://doi.org/10.1016/j.edurev.2016.07.001

Radloff, J. & Guzey, S. (2016). Investigating pre-service STEM teacher conceptions of STEM education. *Journal of Science Education and Technology, 25*, 759-774. https://doi.org/10.1007/s10956-016-9633-5

Vissers, V. (2019). *Energy storage in a world powered by renewables.* [Bachelor Integration project, University of Groningen].

Using Augmented Reality and Simulations to Support Learning of Fundamental Electrical Concepts

Florian Frank, Christoph Stolzenberger, Markus Elsholz and Thomas Trefzger

University of Wuerzburg, Germany

Secondary school students struggle to fully understand fundamental electrical concepts such as the relation and difference between current and voltage. According to multimedia learning theories, the use of augmented reality applications or simulations can support students in this process by helping to manage their cognitive load while learning. A study was conducted with 177 secondary school students, investigating the effects of the use of a simulation and an Augmented Reality (AR)-application for supporting the students while learning about electricity. The use of the simulation and the AR-application was compared with the use of traditional informational graphics for presenting analogy-based models for electricity. Over the course of a student laboratory consisting of four units pertaining to fundamental electrical concepts (electrical current and voltage, electrical resistance, parallel and series circuits), the cognitive load the students perceived was measured. The questionnaire used was examined via a confirmatory factor analysis, after which the data was processed according to an Item-Response-Theory-approach. Subsequent analyses showed that students reported a significantly lower extraneous cognitive load when working with the AR-application as compared to working with the simulation, but no significant differences when compared to working with informational graphics.

Keywords: Physics, multimedia-supported learning, cognitive load.

Introduction

Recent studies of student's misconceptions about electricity have shown that after the introductory lessons on simple electrical circuits, almost two thirds of students have difficulties correctly separating between the concepts of electrical current and voltage (Ivanjek et al., 2021), with more than half the students misunderstanding voltage to be a property of electrical current (Burde, 2018, p. 244). These misconceptions are still prevalent after finishing early secondary school (Müller et al., 2022) and can even be found among first semester university students (Rahmawati et al., 2023). These empirical findings suggest that the way electricity is taught in schools may not be sufficient to enable the students to fully understand electricity and its related core concepts. To enhance

the students' understanding, teachers frequently use analogy models based on systems the students are more familiar with like air pressure (Burde & Wilhelm, 2020). While the use of these models can support learning, the way the models are presented is crucial, since weaknesses in model presentation can lead to further difficulties and inhibit correct understanding (Mogstad & Bungum, 2020)

Based on this analysis, a research project was conducted which aimed to study and thereby gain a more accurate understanding of the difficulties of presenting didactic models for electricity. Traditionally, these models have been taught to the students via the use of informational graphics, which challenged students with the cognitively demanding task of relating the static images to the real-life circuits they are trying to understand. To ease this cognitive burden, we developed digital support material in the form of an Augmented Reality (AR)-application and a simulation, which are designed to present adaptive digital representations of didactic models for electricity fitting precisely to the real-life circuits the students are studying. Findings from Altmeyer et al. (2020) and Thees et al. (2020), relating to the use of AR-applications in learning about electricity, point towards the capacity of AR-applications to improve learning. Altmeyer et al. (2020) report a higher learning gain for AR with equal cognitive load as compared to a non-AR-intervention, whereas Thees et al. (2020) showed an equal learning gain with lower extraneous cognitive load for the AR-intervention. The effects of using the digital support material developed by us on the learning process were evaluated in an experimental study taking place at the student lab. The full research project encompasses the analysis of the student's conceptual knowledge growth, of the cognitive burden placed on the students, and of the time spent on the learning tasks. In this article, we focus on the cognitive load experienced while learning and how it is influenced by the use of multimedia technology (simulations and AR).

Theory and Predictions

Cognitive Load Theory

A central tenet of the *Cognitive Load Theory (CLT)* is that every student's cognitive capacity is finite (Sweller et al. 2019). This cognitive capacity is utilised in learning, where every learning task imposes a certain cognitive load. If that cognitive load exceeds the cognitive capacity, the learning process is hindered. One can distinguish between three categories of cognitive load: intrinsic cognitive load (ICL), extraneous cognitive load (ECL) and germane cognitive load (GCL). Intrinsic load results from processes contributing to an understanding of the learning content and is therefore independent of the learning situation or mode of content presentation. Processes aimed at understanding the presentation of the content result in extraneous load. Germane load is generated

by processing the learning content, storing it in long-term memory and linking it to previously acquired information. Following this, an ideal learning task should primarily impose intrinsic and germane cognitive load, since these are caused by understanding and processing the learning content, and should avoid imposing extraneous load, since it is a product of trying to understand not the learning content, but the presentation thereof.

Cognitive Theory Of Multimedia Learning

The *Cognitive Theory of Multimedia Learning (CTML)*, which adapts these principles of the CLT for complex learning scenarios involving multiple sources of information, contains design guidelines that aim to reduce the overall cognitive load imposed by the learning situation (Mayer, 2014). Following the principles of segmentation and coherence, the learning content should be separated into the small units, containing only the necessary information, thereby reducing the intrinsic load imposed to the minimum that is necessary (Mayer & Moreno, 2010). Regarding the extraneous load, according to the principles of spatial and temporal contiguity, the ECL can be reduced by making all resources relevant to the learning process available at the same time and in the same place, thereby reducing the need for taxing use of short-term memory storage (Mayer & Fiorella, 2014).

Key Considerations And Hypotheses

We propose that the central resources relevant to a learning process regarding electricity are the real-life experiment (e.g. a simple circuit with a light bulb) and the didactic model based on an analogy to understand the mechanisms imperceptible in the experiment (e.g. a model based on a height analogy explaining how the battery causes an electrical current, also see Mogstad & Bungum, 2020). Since the resources are partly real/observable (experiment) and partly theoretical/imaginary (the didactic model), the presentation of both at the same time and in the same place is not possible while using informational graphics to convey the didactic model. However, a simulation might be used to superimpose the model atop a digital replica of the circuit, thereby combining both resources, albeit in a solely digital environment. The use of a see-through-AR-application allows the presentation of the model atop the real-life circuit, theoretically strengthening the connection between the experiment and the model even further.

We designed a student lab program consisting of four units of increasing complexity and internal difficulty to study the effects of the use of a simulation and an AR-application on the perceived cognitive load compared to using informational graphics to present the didactic models. Our hypotheses were the following:

- H1: The intrinsic cognitive load does not differ between the three modes

of model presentation.

- H2: The intrinsic cognitive load differs between the four units.
- H3: The extraneous cognitive load differs between the three modes of model presentation.
- H4: The germane cognitive load differs between the three modes of model presentation.

Methods

We used an experimental design with multiple tests to compare a traditional approach (intervention group IG) to present didactics models with a simulation-based approach (interventions group SIM) and an AR-based approach (intervention group AR) at four different complexity levels.

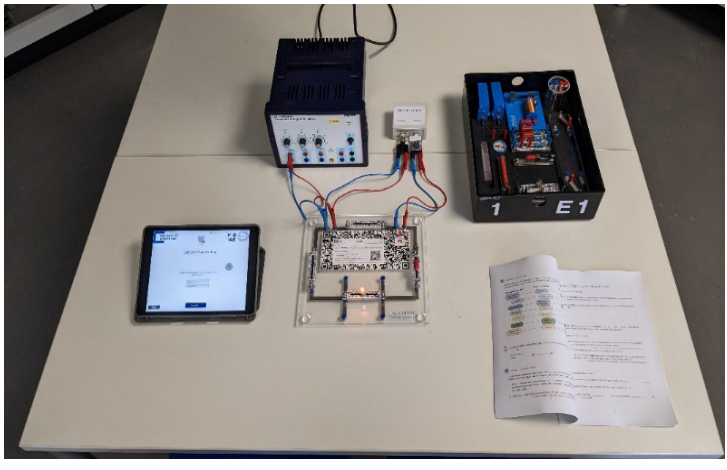
Participants

The investigations were carried out in 2022 and 2023. The participants attended local eight-grade secondary school classes. Before the students took part in the intervention, they had already completed their lessons on electricity and its basic concepts (electrical current, voltage, and resistance). In total, N = 177 8th-grade students participated and were randomly assigned to an intervention group (N = 67 in group IG, N = 64 in group SIM, and N = 46 in group AR).

Intervention Design

The intervention took place in an out-of-school student laboratory at the University of Würzburg. In the student lab, the participants completed four units in a fixed order, starting with a unit on "Electrical Current and Voltage". This unit was designed to primarily contain previously taught information to give the students an opportunity to familiarize with the student lab and the digital support materials they would be working with. The next unit, "Electrical Resistance", involved new content and unfamiliar tasks, e.g. using the models presented in the digital support material to explain simple observations. Units 3, "Parallel Circuits", and 4, "Serial Circuits", required the participants to use the models presented in the digital support material to form hypotheses about Kirchhoff's circuit laws, involving complex reasoning and justification tasks. During units 1 to 4, the students used two didactic models of electricity (based on a height analogy of a marble track and based on an air pressure analogy, respectively) to learn about electricity.

Figure 1. Material used in the student lab (from left to right: tablet computer hosting the digital support materials, experimentation set for simple circuits, printed-out workbook)



The students worked in small groups, supervised by university student teachers. The students were supplied with a printed-out workbook, containing the instructions and tasks described above, as well as an experimentation set for simple circuits and a tablet computer (see Figure 1). The structure and contents of the workbook are the same across all intervention groups. At multiple points in the workbook, the students are instructed to use the tablet to access presentations of the didactic models. When developing the intervention, care was taken to ensure that all groups were asked to interact with the digital support material at the same time to equalize the motivational effects of technology use across all groups as much as possible. After each unit, the students were asked to fill out a short test, followed by a 10–15-minute rest before starting with the next unit.

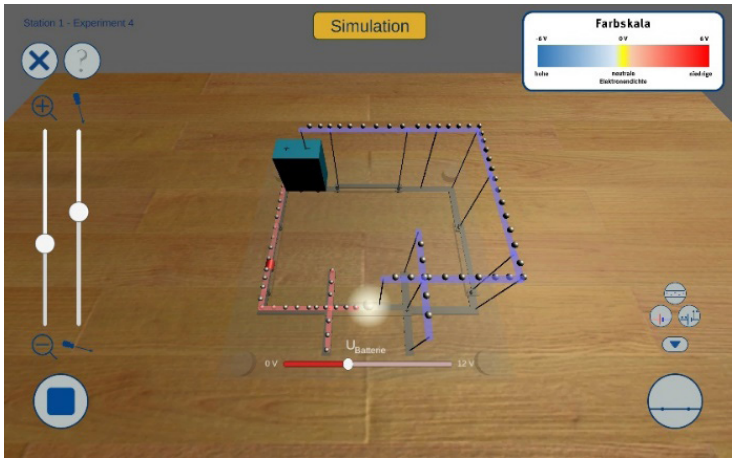
Differences Between Intervention Groups

The traditional visualization of the models via infographics (intervention group IG) was provided to the students via the online platform *tet.folio* (Haase, Kirstein & Nordmeier, 2016). To that end, a tet.book was created that contained informational graphics taken from ordinary schoolbooks used in Germany. The graphics were in part supported by short annotations. The goal of the tet.book called "Visualisations" was to present the didactic models in a concise and familiar way.

To present the didactic models for intervention group SIM, a simulation was developed, that contained a three-dimensional replica of the real experimentation set used in the student lab. The students recreate the real-life circuit they had built via the simulation by dragging the respective components to their spot. After the

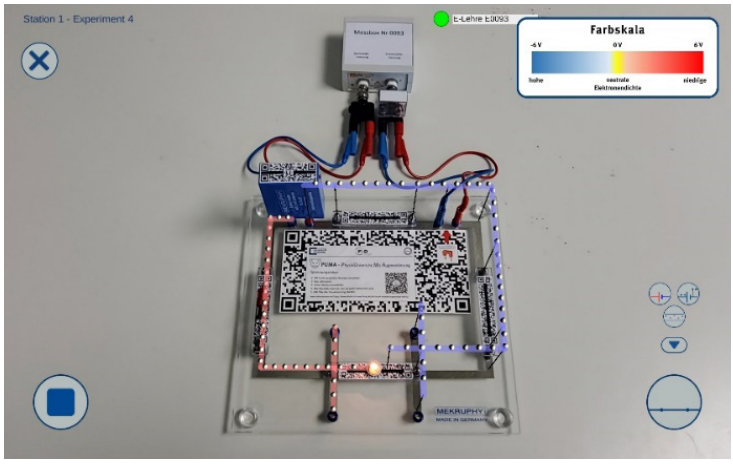
construction of the virtual circuit is completed, the students use the simulation to visualize the didactic models, which were fitted to the digital circuit and adapted to changes in the digital circuit (e.g. when closing or opening a switch) in real time (see Figure 2).

Figure 2. Simulation-based approach to presenting the didactic models for electricity (shown are an air-pressure- and a height-based model in the context of a simple circuit with one light bulb)



For intervention group AR, a similar application has been developed, using augmented-reality-technology to overlay the didactic models atop the real-life circuit (Stolzenberger et al., 2022). It works in the same way as the simulation described above: after building a circuit, you use the camera of the tablet computer to scan the experiment. On successful scan, the application then visualizes the didactic models by superimposing digital representations of them onto the real-time camera feed of the tablet computer (see Figure 3). Using a small Bluetooth-measurement box, a connection between the circuit and the AR-application is established, which again allows for changes made to the circuit to impact the digital representation of the didactic models.

Figure 3. AR-based approach to presenting the didactic models for electricity (shown are an air-pressure- and a height-based model in the context of a simple circuit with one light bulb)



Dependent Variables And Test Instruments

Cognitive Load

The cognitive load was measured using the Cognitive Load Scale (Klepsch et al., 2017). The scale contains 8 items, 2 pertaining to measuring the intrinsic load, 3 measuring the extraneous load, and 3 measuring the germane load. The authors themselves however suggest only using 2 of the 3 germane load-items, since one item is "only useful if GCL is varied on purpose in a given learning material (e.g., through prompts)" (Klepsch et al., 2017, p. 10). Every item contains a statement about the learning task, the students are asked to note their agreement or disagreement on a 7-point Likert scale.

The scale in its entirety (all 8 items) was administered during the short post-test after every unit, in total 4 times.

Conceptual Knowledge and Time on Task

The conceptual knowledge development was measured using the 2T-SEC Test developed by Ivanjek et al. (2021). It was administered in the pre-test, together with the measurement of the academic achievement of the students (self-reported grades in mathematics and science), and in the four post-unit-tests, together with the cognitive load scale. For the post-tests, the original test was divided evenly into four parts pertaining to the contents of the four units.

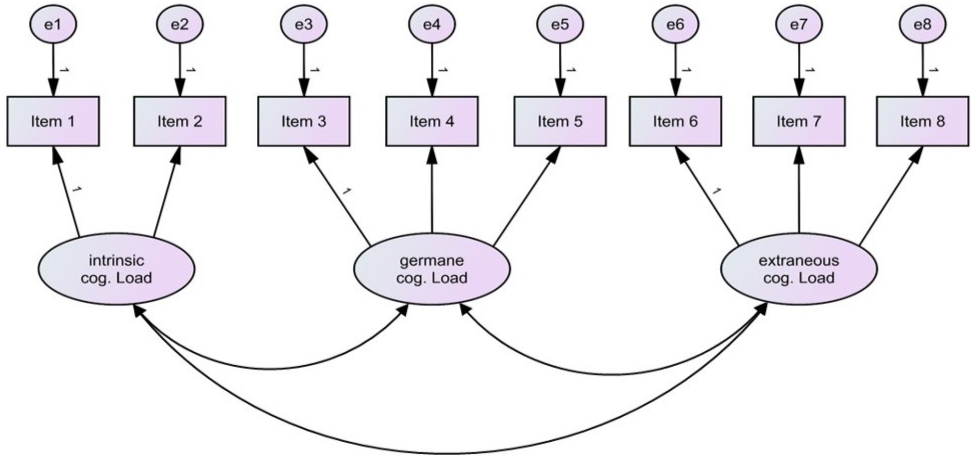
The Time-on-Task was logged by the teaching assistants, totaling to four separate Time-on-Task-measurements per every student.

Results

Preparing The Cognitive Load Data

We started by examining the structure of the Cognitive Load Scale. Although the material used by the intervention groups did not differ in the number of prompts or similar didactic scaffolding, we purposefully used the full 8-item scale with the intent of using a confirmatory factor analysis to support the authors claim of leaving out one item pertaining to the germane load.

Figure 4. Structure of the Cognitive Load Scale (full 8-item-variation, the 7-item-variation removes item 5)



SPSS AMOS was used to model the structure (see Figure 4) and compare the two versions of the instrument. The structure was confirmed for each post-unit-test separately, so four times in total. The model fit was evaluated using the probability of obtaining the χ^2-Value (higher χ^2-P indicates better fit), the comparative fit index (higher CFI indicates better fit), the root-mean-square error of approximation (lower RMSEA indicates better fit) and the standardized root mean squared residual (lower SRMR indicates better fit). Following Hu & Bentler (1999), an acceptable fit would result in the CFI exceeding .95, with the RMSEA and the SRMR staying below .06 and .08, respectively. Comparing the model fit indices obtained by our analysis of the cognitive load scale, we concluded the model containing 7 items to be the better fitting model (see Table 1 for details). Although the 7-item-variation fits our data better, it must be noted, it did not result in only passable fit indices. While the SRMR showed acceptable values, only half of the past-unit-models achieved a CFI of above .95 and none

reached a RMSEA-value of below .06. Klepsch et al. (2017) analyzed the 7-item-variation as well and reported a χ^2-P-value of below .001, a CFI-value of .97 and a RMSEA-value of .021. Further analysis will be conducted to investigate this discrepancy between their reported fit indices and the fit indices found for our data.

Table 1. Model Fit of the two Cognitive Load Scale Variations (the SRMR-value of the 7-item-variation for Unit 1 could not be computed by SPSS AMOS).

	8-item-variation				7-item-variation			
	χ^2-P	CFI	RMSEA	SRMR	χ^2-P	CFI	RMSEA	SRMR
Unit 1	< .001	.829	.121	.0925	< .001	.908	.101	*
Unit 2	< .001	.903	.109	.0757	< .001	.921	.113	.0675
Unit 3	< .001	.920	.110	.0802	.005	.969	.077	.0492
Unit 4	< .001	.913	.124	.0846	< .001	.955	.100	.0686

To prepare the Cognitive Load Data for further analysis, the R-Package TAM (Robitzsch et al., 2022) was used to process the raw data and scale it according to an Item-Response-Theory-Approach (Embretson & Reise, 2009, p. 40-61). Since the Cognitive Load Scale splits into three separate subscales pertaining to each of the types of cognitive load, we used three separate IRT-models to process the data. Each of these models contained the repeated measures Likert-data for one of the cognitive load types, resulting in a four-dimensional (for the four post-unit tests) Generalized Partial Credit (to process the Likert-data) model (Embretson & Reise, 2009, p. 110-115). Preparing the data this way results in factor scores representing the perceived cognitive load, where the higher the factor score, the higher the perceived cognitive load.

Analysis Of Group-Wise Intrinsic Load Data

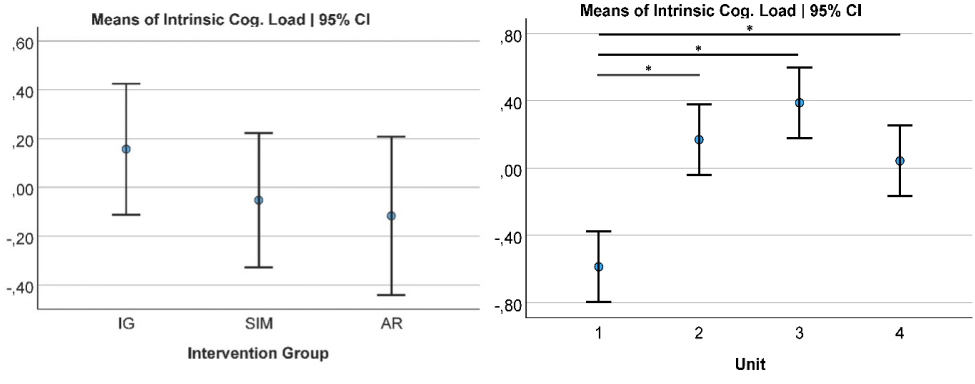
To interpret the differences between the intervention groups, we calculated the mean intrinsic cognitive load experienced across the four units. Afterwards, a One-Way-ANOVA was conducted to investigate the effect of the model presentation on this mean intrinsic cognitive load. We found no statistically significant main effect of intervention group (and therefore model presentation) on the mean intrinsic cognitive load, $F(2,174)=.980$, $p=.377$, partial $\eta^2=.011$ (see Figure 5).

Analysis Of Unit-Wise Intrinsic Load Data

We conducted a One-Way-ANOVA to investigate the effect of the content learned during the different units on the experienced intrinsic cognitive load. We found a statistically significant main effect of the unit (and therefore the learning content) on the intrinsic cognitive load, $F(3,700)=15.409$, $p<.001$,

partial η^2=.062. Post-Hoc Tests indicated that the mean intrinsic cognitive load in unit 1 (M=-.586, SD=.107) was significantly different (p<.05) from the mean intrinsic cognitive load in unit 2 (M=.169, SD=.107), unit 3 (M=.388, SD=.107) and unit 4 (M=.043, SD=.107). No other statistically significant differences in intrinsic cognitive load were found (see Figure 6).

Figure 5 & 6. Means of the Intrinsic Cognitive Load by model presentation (on the left) and by unit (on the right)



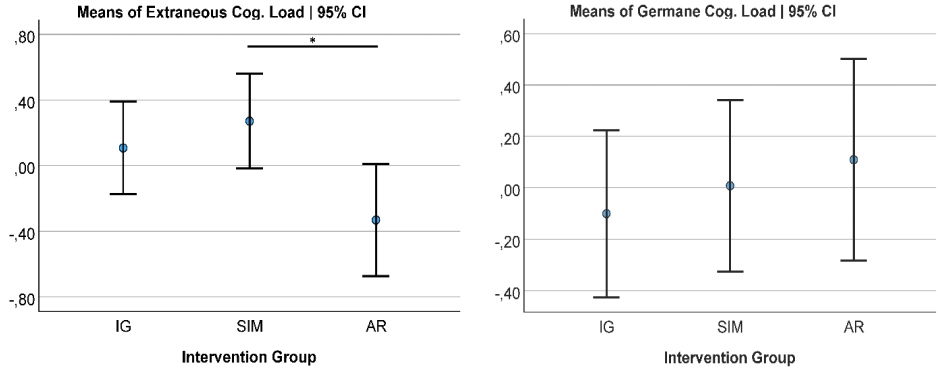
Analysis Of Group-Wise Extraneous Load Data

To interpret the differences between the intervention groups, we calculated the mean extraneous cognitive load experienced across the four units. Afterwards, a One-Way-ANOVA was conducted to investigate the effect of the model presentation on this mean extraneous cognitive load. We found a statistically significant main effect of intervention group (and therefore model presentation) on the mean extraneous cognitive load, F(2,174)=3.663, p=.028, partial η^2=.040. Post-Hoc Tests indicated that the mean extraneous cognitive load in intervention group AR (M=-.332, SD=.173) was significantly different (p<.05) from the mean extraneous cognitive load in intervention group SIM (M=.271, SD=.146). The mean extraneous cognitive load in intervention group IG (M=.108, SD=.143) did not significantly differ from the mean extraneous cognitive load in intervention groups SIM (p=.811) and AR (p=.147) (see Figure 7).

Analysis Of Group-Wise Germane Load Data

To interpret the differences between the intervention groups, we again calculated the mean germane cognitive load experienced across the four units. Afterwards, a One-Way-ANOVA was conducted to investigate the effect of the model presentation on this mean germane cognitive load. We found no statistically significant main effect of intervention group (and therefore model presentation) on the mean germane cognitive load, F(2,174)=.336, p=.715, partial η^2=.004 (see Figure 8).

Figure 7 & 8. Means of the Extraneous Cognitive Load by model presentation (on the left) and the Germane Cognitive Load by model presentation (on the right)



Discussion and Conclusion

Regarding The Intrinsic Cognitive Load

The intrinsic cognitive load did not differ between the modes of model presentation, hypothesis H1 is therefore accepted. This finding supports our claim that the content the students learned during the intervention did not differ between the different modes of model presentation. Great care was taken when designing the intervention to ensure this being the case, since the goal of our research was to investigate different modes of presenting and learning the same content.

The unit-wise analysis revealed a difference in intrinsic cognitive load between unit 1 and units 2, 3 & 4, hypothesis H2 is therefore accepted. As previously stated, Unit 1 focused on familiarizing the participants with the student lab and the material, therefore primarily containing already taught content. Units 2 to 4 differed from this and confronted the students with new content and tasks. This finding has been described as the Pre-Training-Principle, which states that when learning, the intrinsic cognitive load can be reduced by using previously stored knowledge to help chunk the incoming information (Mayer & Moreno, 2010, p. 145-146.

Regarding The Extraneous Cognitive Load

It was hypothesized that the use of multimedia for presenting the models would result in a difference in extraneous cognitive load. Although a significant difference between the modes of model presentation was found, leading us to accept hypothesis H3, it was expected to be a significant reduction when using the simulation and the AR-application. Instead, the AR-application induced significantly lower extraneous cognitive load as compared to the simulation,

while the use of traditional informational graphics did not significantly differ from these two modes of presentation. We propose the following interpretation of this finding: The central learning resources are, as stated previously, the real-life experiment and the intangible didactic model. The AR-application manages to combine both by superimposing a representation of the didactic model atop the real-life experiment. The simulation tries to achieve a combination of both resources by substituting the real-life experiment with a digital one, thereby introducing a new source of content presentation not present in the AR-application. This likely lead to the students having to process more simultaneous content presentation, thereby raising their perceived extraneous cognitive load. Although we found a difference in mean extraneous cognitive load between intervention groups IG and AR, it was not statistically significant.

Regarding The Germane Cognitive Load

The analysis revealed there to be no significant difference in mean germane cognitive load between the three modes of model presentation, hypotheses H4 is therefore rejected. Initially, the alternative hypothesis was chosen, because we presumed that learning about electricity overburdens the students cognitively. If the extraneous cognitive load while learning could be reduced, capacity would be freed up to be used for processing and memorizing the learning content, which would result in a higher germane cognitive load. Since the different modes of model presentation were presumed to yield different extraneous cognitive loads, one would then also expect a difference in germane cognitive load. Because we found no difference in germane cognitive load, we conclude that at least over the course of the student lab, we did not cognitively overburden the participants.

Conclusion

In an experimental study taking place at an out-of-school student laboratory, research was conducted to investigate the effect the use of multimedia technology for presenting analogy-based models for electricity (in the form of a simulation and an AR-application) has on the cognitive burden induced by the learning process. It was found that the use of an AR-application can significantly reduce the extraneous cognitive load caused by the processing of the content presentation as compared to the use of a simulation. Using informational graphics to present the models resulted in a non-significant difference in extraneous cognitive load compared to the use of the simulation and the AR-application. Our cumulative findings affirm known principles of the *Cognitive Theory of Multimedia Learning*, more precisely the Principles of Temporal and Spatial Contiguity (pertaining to the extraneous cognitive load, shown by the reduction of the load when using the AR-application, Mayer & Fiorella, 2014) and the Pre-Training-Principle (pertaining to the intrinsic cognitive load, Mayer &

Moreno, 2010). Meta-Reviews by Radu (2014) and Buchner et al. (2022) report the use of AR-application in learning to have a positive effect on knowledge gain, group collaboration, and perceived cognitive load when compared to non-AR-learning. Radu (2014) also reports that AR-applications can overburden learners if implemented non-optimally, which is akin to our finding for the use of a simulation for model presentation.

The results presented here are preliminary, the study is currently ongoing and will in total include approximately N = 360 participants. Future analyses will also include the data on the conceptual knowledge development and Time-on-Task, as well as constructs presumed to be moderating variables (i.e. the technological affinity or the visualization capabilities of the participants).

The materials used in the presented research (including the AR-application and the simulation) are made available for free at https://go.uniwue.de/puma-s. Additionally, the multimedia applications have been published in the Apple App-Store and the Google Play Store under the name "PUMA : Spannungslabor".

Acknowledgement

This project is part of the "Qualitätsoffensive Lehrerbildung", a joint initiative of the Federal Government and the Länder which aims to improve the quality of teacher training. The programme is funded by the Federal Ministry of Education and Research. The authors are responsible for the content of this publication.

References

Altmeyer, K., Kapp, S., Thees, M., Malone, S., Kuhn, J., & Brünken, R. (2020). The use of augmented reality to foster conceptual knowledge acquisition in STEM laboratory courses—Theoretical background and empirical results. *British Journal of Educational Technology*, *51*(3), 611–628. https://doi.org/10.1111/bjet.12900

Buchner, J., Buntins, K., & Kerres, M. (2022). The impact of augmented reality on cognitive load and performance: A systematic review. *Journal of Computer Assisted Learning*, *38*(1), 285–303. https://doi.org/10.1111/jcal.12617

Burde, J.-P. (2018). *Konzeption und Evaluation eines Unterrichtskonzepts zu einfachen Stromkreisen auf Basis des Elektronengasmodells*. Logos Verlag Berlin.

Burde, J.-P., & Wilhelm, T. (2020). Teaching electric circuits with a focus on potential differences. *Physical Review Physics Education Research*, *16*(2), 020153. https://doi.org/10.1103/PhysRevPhysEducRes.16.020153

Embretson, S. E., & Reise, S. P. (2009). *Item response theory for psychologists* (Reprinted 2009 by Psychology Press). Psychology Press.

Haase, S., Kirstein, J., & Nordmeier, V. (2016). The Technology Enhanced Textbook: An HTML5-based Online System for Authors, Teachers and Learners. In L.-

J. Thoms & R. Girwidz (Eds.), *Selected Papers from the 20th International Conference on Multimedia in Physics Teaching and Learning* (pp. 85-92). Mulhouse: European Physics Society.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Ivanjek, L., Morris, L., Schubatzky, T., Hopf, M., Burde, J.-P., Haagen-Schützenhöfer, C., Dopatka, L., Spatz, V., & Wilhelm, T. (2021). Development of a two-tier instrument on simple electric circuits. *Physical Review Physics Education Research*, *17*(2), 020123. https://doi.org/10.1103/PhysRevPhysEducRes.17.020123

Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and Validation of Two Instruments Measuring Intrinsic, Extraneous, and Germane Cognitive Load. *Frontiers in Psychology*, *8*, 1997. https://doi.org/10.3389/fpsyg.2017.01997

Mayer, R. E. (2014). Cognitive Theory of Multimedia Learning. In R. E. Mayer (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (2. Aufl., S. 43–71). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.005

Mayer, R. E., & Fiorella, L. (2014). Principles for Reducing Extraneous Processing in Multimedia Learning: Coherence, Signaling, Redundancy, Spatial Contiguity, and Temporal Contiguity Principles. In R. E. Mayer (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (2. Aufl., S. 279–315). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.015

Mayer, R. E., & Moreno, R. (2010). Techniques That Reduce Extraneous Cognitive Load and Manage Intrinsic Cognitive Load during Multimedia Learning. In J. L. Plass, R. Moreno, & R. Brünken (Hrsg.), *Cognitive Load Theory* (1. Aufl., S. 131–152). Cambridge University Press. https://doi.org/10.1017/CBO9780511844744.009

Mogstad, E., & Bungum, B. (2020). Ski lifts, bowling balls, pipe system or waterfall? Lower secondary students' understanding of analogies for electric circuits. *Nordic Studies in Science Education*, *16*(1), 37–51. https://doi.org/10.5617/nordina.6882

Radu, I. (2014). Augmented reality in education: A meta-review and cross-media analysis. *Personal and Ubiquitous Computing*, *18*(6), 1533–1543. https://doi.org/10.1007/s00779-013-0747-y

Rahmawati, R., Widiasih, W., Marisda, D. H., & Riskawati, R. (2023). Using Four-Tier Test to Identify Prospective Elementary Teacher Students' Misconception on Electricity Topic. *Jurnal Penelitian Pendidikan IPA*, *9*(10), 7793–7802. https://doi.org/10.29303/jppipa.v9i10.3272

Robitzsch, A., Kiefer, T., & Wu, M. (2022). *TAM: Test Analysis Modules*. https://CRAN.R-project.org/package=TAM

Stolzenberger, C., Frank, F., & Trefzger, T. (2022). Experiments for students with built-in theory: 'PUMA: Spannungslabor' – an augmented reality app for studying electricity. *Physics Education*, *57*(4), 045024. https://doi.org/10.1088/1361-6552/ac60ae

Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Thees, M., Kapp, S., Strzys, M. P., Beil, F., Lukowicz, P., & Kuhn, J. (2020). Effects of augmented reality on learning and cognitive load in university physics laboratory courses. *Computers in Human Behavior*, *108*, 106316. https://doi.org/10.1016/j.chb.2020.106316

Wilhelm, T., Müller, S., & Burde, J.-P. (2015). Vergleich von Schülervorstellungen zur Elektrizitätslehre in Hessen und Weißrussland. *PhyDid B - Didaktik der Physik - Beiträge zur DPG-Frühjahrstagung*, *0*(0). https://ojs.dpg-physik.de/index.php/phydid-b/article/view/629



Strand 14

Evaluation and Assessment in Teaching and Learning

Jens Dolin & Lukas Rokos

Strand Chairs/Co-editors

Foreword

This chapter, titled Evaluation and Assessment in Teaching and Learning, forms an important part of the ESERA 2023 Conference Proceedings. It brings together seven studies that focus on the critical role of evaluation and assessment in shaping effective science education. In the evolving landscape of teaching and learning, assessment serves as a key mechanism for understanding student progress, refining instructional practices, and driving educational improvement.

The proceedings in this chapter explore a wide range of approaches to assessment, from traditional methods to more innovative, formative techniques designed to support student learning in dynamic ways. The studies highlight how effective evaluation practices can provide meaningful feedback, promote student engagement, and enhance conceptual understanding. They also examine the challenges involved in designing assessments that are fair, reliable, and aligned with educational goals.

In addition to assessing students' knowledge and skills, the research in this chapter considers the broader implications of assessment on teaching strategies, curriculum design, and educational equity. Topics such as the role of self-assessment, peer feedback, and data-driven decision-making are explored, offering practical insights for educators aiming to create more responsive and supportive learning environments.

The findings presented in this chapter underscore the importance of viewing assessment not merely as a tool for measuring outcomes but as an integral part of the learning process. By focusing on both formative and summative approaches, these studies emphasize the value of assessments that guide and enhance the learning journey, providing opportunities for reflection, growth, and continuous improvement.

We are grateful to the authors and reviewers whose contributions have made this chapter possible. Their work adds significant value to the ongoing discussions around best practices in evaluation and assessment within science education. As you engage with the research presented here, we hope that you will find ideas and strategies that resonate with your own teaching and research endeavors.

This chapter serves as a resource for educators, researchers, and policymakers dedicated to refining how we assess learning and how we can use evaluation as a catalyst for better educational outcomes. We trust that the insights shared will inspire further exploration and innovation in the field of science education assessment.

Development of a Scientific Practices Survey for Middle School Students

Yann Shiou Ong, Yew-Jin Lee, Luke Gek Siong Liau, Xu Kang Tan

National Institute of Education, Nanyang Technological University, Singapore

The new Singapore science curriculum framework has been revised to incorporate an important domain—the Practices of Science—which closely aligns with the US Next Generation Science Standards' scientific and engineering practices in the local form of Ways of Thinking and Doing in Science (WOTD). Given a lack of appropriate survey instruments for evaluating students' grasp of scientific practices for large groups of learners in a pragmatic way, we set out to develop such an instrument, in anticipation of research and assessment needs. A scientific practices survey instrument (version 3) was thus constructed and validated in a pilot study with middle school students (Grades 7-8) in a Singapore secondary school (n=82). Our items were lean in science content, but scientific practices rich and inspired by scientific investigations reported in science journals/popular science websites. They mapped onto practices in the three spheres of scientific activity: investigating, evaluating, and explaining. All items were validated and achieved reasonable inter-rater reliability based on Cohen's kappa measure. Rubrics for each item were written to reflect four performance levels: extending (highest), proficient, developing, and emerging (lowest). Pilot study findings suggest most students were operating at the developing level: Many students did not demonstrate a grasp of the scientific practices, but instead, followed typical "school science" practices.

Keywords: Assessment, Scientific Practices, Survey instrument

Need For a Scientific Practices Survey Instrument

The Singapore science curriculum framework has recently been revised to incorporate an important new domain–the Practices of Science (POS) –for Grades 3-12 science curricula. Based on the Singapore Ministry of Education's conceptual framing, POS includes three dimensions: (a) eight aspects of ways of thinking and doing in science (WOTD) that closely resemble the USA Next Generation Science Standards' scientific and engineering practices and fall under the three spheres of scientific activity (i.e. investigating, developing explanations, evaluating [Osborne, 2014]), (b) the nature of scientific knowledge (NOS), and (c) connections between science, technology, society, and the environment (STSE).

Our study focuses on students' demonstrated scientific practices that lies at the heart of thinking critically and reflectively about science, and resembles what scientists do. To determine students' grasp of scientific practices, robust and validated assessment instruments are helpful although limited in the literature. Researcher-made survey instruments for scientific inquiry (e.g., Views about Scientific Inquiry (VASI) (Lederman et al., 2014) and scientific argumentation (Osborne et al., 2016) are available although there are no validated assessments of students' grasp of a range of WOTD/scientific practices through pen-and-paper instruments. While in-depth assessments of individual scientific practices do exist e.g., through video analysis of students' scientific argumentation (Sampson et al., 2011)called Argument-Driven Inquiry (ADI or scientific modelling (Manz, 2012)there is relatively little research on how learning environments can support the simultaneous, coordinated development of both practice and the knowledge that emerges from and supports scientific activity. This study reports on the co-construction of modeling practice and ecological knowledge by following the development of one seminal disciplinary concept, plant reproduction, through third graders' yearlong investigation of a wild backyard area. Representational activity facilitated the visibility and utility of meanings for reproduction; these meanings, in turn, shaped students' subsequent modeling practice. Over time, shifts were evident in both the community's meanings for reproduction and their framing of meanings in relation to modeling activity. The analysis affirms the utility of attending to student knowledge as it is used in practice to navigate complexity and uncertainty, rather than as assimilation of declarative statements. It also provides images of productive relations between modeling and knowledge, with implications for how the two might bootstrap each other in extended classroom investigation. (c, such methods involving practical activities are very challenging for assessing large numbers of learners quickly. Our study aims to address these gaps in scientific practices assessment by constructing and validating a new written survey instrument to assess middle school students' scientific practices within the three spheres of scientific activity. We chose these grade levels as the POS was most clearly articulated in the official middle school science curriculum document implemented in 2021. Our new *scientific practices survey instrument* potentially helps to determine (i) secondary students' performance levels on scientific practices in the respective spheres of scientific activity, and (ii) potential changes in students' performance levels due to educational interventions.

Scientific Practices Survey Development

We initially constructed versions A and B of the scientific practices survey to test more items while ensuring each could be completed within 45-60 minutes by middle school students without experiencing survey fatigue. Based on the pilot study findings, we decided to focus on developing version B. Hence, this

study reports only on the findings relevant to version B (henceforth, scientific practices survey version 3).

Survey items were lean on science content knowledge (appropriate for seventh graders) and instead, focused on scientific practices as intended. Items were inspired by scientific investigations reported in science journals or popular science websites then mapped onto the three spheres of scientific activity. The mapping of items is shown in Table 1.

Table 1. Items distribution across spheres of activity in survey.

Scientific activity spheres and scientific practices (Q1) Pond Skater		Scientific Practices Survey (version 3) questions			Items per practice	Items per sphere
		(Q2) Spider Web	(Q3) Lizard			
Investigating	Asking questions		a(i), a(ii) [analyzed together]		2	**5**
	Planning and carrying out investigations	a(i), a(ii) [analyzed together]	b(ii)		2	
	Analyzing and interpreting data		b(i), b(iii)		1	
Evaluating	Engaging in argument from evidence	b(i), b(ii) [together]	(c)		2	**2**
Explaining	Developing and using models	(c)		b(i), b(ii), b(iii)	3	**5**
	Constructing explanations			(a)	2	

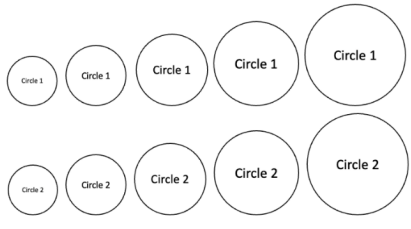
For brevity, part of one question (Pondskater) is presented in Figure 1.

Figure 1. One item in the Pondskater question.

Q1. Janice sees some Pond Skaters in the school garden. She thinks that since a Pond Skater spreads its legs wide across the water (see Figure 1a), this lets it spread its weight. Thus, each leg has less pressure on the water and it does not break through the water surface. This allows Pond Skaters to float on the water surface [Image of actual pondskaters on water not reproduced here].

Janice wanted to investigate if spreading weight over a larger area helps things to float. She came up with the following hypothesis: **Objects with a larger surface area will float more often than objects of the same mass with a smaller surface area.**

Janice then modelled Pond Skaters using wires. She cut 10 pieces of wire of different lengths from the same coil of wire. The pieces of wire are shown in the figure below.



Next, Janice twisted the wires to make circles of 5 different sizes - there are two circles for each size. She placed these circles in a tray of water and observed if they float or sink. Her data are shown in the table below.

Circle area / cm²	**Circle 1**	**Circle 2**
3.50	Sink	Sink
6.80	Sink	Float
11.00	Float	Sink
16.30	Float	Float
25.50	Float	Float

1(a)(i). Based on the data, how many more times should Janice repeat her investigation with more circles of each area?

Select your answer from the following options.

A. Janice does not need to repeat her investigation.

B. Only once more, so there will be three sets of data for the same circle area: Circle 1, Circle 2, and Circle 3.

C. As many times as possible.

D. None of the above.

( )

1(a)(ii). Explain your answer to 1(a)(i).

A rubric for each item was written to reflect performance at four levels: extending (highest), proficient, developing, and emerging (lowest). Table 2 shows the rubric and example responses for the item in Figure 1.

Table 2. Rubric and example responses for Q1a(i) and Q1a(ii).

Level	Description	Examples of Student Responses
Extending	Justification improves the reliability of the investigation results with reference to the given data to justify the need for repeat and mentions what counts as sufficient to stop repeating (i.e., stop criteria).	*C. There is a difference in results for circles of areas 6.80 and 11.00 cm^2. Thus, it would be better to repeat the experiment as many times as possible to see which result is most likely to happen, and to check if the results for other circle areas are consistent.*
Proficient	Justification improves the reliability of the investigation results with reference to the given data to justify the need to repeat but does not mention what counts as sufficient to stop repeating.	*B. We can find the most reliable answer which appeared 2 out of 3 times or 3 out of 3 times in the experiment.* [Note: 2 out of 3 times for the 6.80 and 11.00 cm^2 circle areas is not a clear majority,]
Developing	Justification considers reliability but is generic and can be applied to any investigation.	B. Repeat once more to find an average so that the results would be reliable. B. 3 sets of data is more accurate/ reliable
Emerging	Justification does not consider reliability OR response does not include justification for repeating the experiment OR response does not make sense or there is no justification.	*B. Janice should repeat one more time as a minimal of 3 trials are needed.* [Note: Response recognises the need to repeat the experiment but did not indicate the reason for repeating the experiment]

Pilot Study Context

A total of 82 eighth graders from a high-performing secondary school in Singapore completed the scientific practices survey (version 3) in the pilot study. Permission was sought and obtained from students and their parents/guardians to participate in the study (IRB no: IRB-2021-669). Students were recommended to complete the online survey within an hour independently and unsupervised at home. The survey setting permitted students to continue after long intermissions and modify their responses before submission.

Validity and Reliability

Survey items demonstrate face and construct validity as determined by the first two authors who are experts in science education experienced in scientific practices. Also, the items have stood up to peer critique at presentations to MOE curriculum developers familiar with POS and international audiences of science education researchers and practitioners. Item rubrics reliability have been tested by two research team members: the third author/a research assistant and the first/ second author independently coded 20 responses from the pilot study. Cohen's

kappa for the items was determined with items reaching strong/almost perfect (0.81–1.00 for seven of 12 items) or moderate agreement (0.66–0.79 for five items).

Findings and Discussion

The pilot study yielded interesting findings, though for brevity, we present only the findings for the Pondskater question in Table 3. Students performed similarly across the Pondskater question items targeting each of the three spheres of scientific activity, with the majority at the developing level.

Table 3. Students' performances across items in the Pondskater question.

Scientific practices (Q1) Pond Skater		**Findings**		**Responses indicating inadequate grasp of practice**
		n = 82 [Level: nos. (by %)]		
Investigating	Planning and carrying out investigations	a(i), a(ii) [together] *Repeat experiment? How many more times?*	Extending: 2 (2.4%) Proficient: 9 (11.0%) **Developing: 56 (68.3%)** Emerging: 15 (18.3%)	Standard/Generic responses e.g., repeat to improve accuracy/ reliability. 3 sets of data. Find "average".
Evaluating	Engaging in argument from evidence	b(i), b(ii) [together] *Accept hypothesis? Why might someone disagree?*	Extending: 0 (0%) Proficient: 4 (4.9%) **Developing: 50 (61.0%)** Emerging: 28 (34.1%)	Most agreed with hypothesis despite recognising problem with data or experiment. Standard/Generic responses e.g., small sample size; not enough repeats.
Explaining	Developing and using models	(c) *How to make a better model of pond skater for experiment?*	Extending: 1 (1.2%) Proficient: 5 (6.1%) **Developing: 56 (68.3%)** Emerging: 20 (24.4%)	Could not recognize critical parts of pond skater to model in experiment. Did not distinguish between model and reality.

Overall, responses indicated many students did not demonstrate a grasp of scientific practices but instead, followed typical "school science" practices, such as "finding average" of categorical float/sink data, agreeing with given hypothesis despite recognizing a problem with the data or method of experimentation, providing generic responses on improving reliability/accuracy (e.g., increase sample size), or describing processes for constructing a physical model without highlighting essential features of the model.

Conclusion/Future Work

The current scientific practices survey will be further validated by performing Rasch analysis to consider item difficulty and determine if/which items should be modified to ensure the survey instrument is sufficiently robust to function as a useful ruler for demonstrating a range of student performances. The revised survey instrument will be tested with students with a range of anticipated performance levels from various schools.

Acknowledgement

We would like to thank our former undergraduate research student, Mr. Yi Jie Soh, for working on previous versions of the survey, as well as our Research Assistant, Miss Miechie, for supporting part of the analysis. We would also like to thank the students for participating in the study and their teachers for assisting with relevant administrative matters.

Funding Statement

Part of the analysis of this study was funded by Singapore Ministry of Education (MOE) under the Education Research Funding Programme (OER 20/22 OYS) and administered by National Institute of Education (NIE), Nanyang Technological University, Singapore. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Singapore MOE and NIE.

Survey Availability

Interested readers may write to the first author at yannshiou.ong@nie.edu.sg to request for the scientific practices survey (version 3).

References

Lederman, J. S., Lederman, N. G., Bartos, S. A., Bartels, S. L., Meyer, A. A., & Schwartz, R. S. (2014). Meaningful assessment of learners' understandings about scientific inquiry - The views about scientific inquiry (VASI) questionnaire. Journal of Research in Science Teaching, 51(1), 65–83. https://doi.org/10.1002/tea.21125

Manz, E. (2012). Understanding the codevelopment of modeling practice and ecological knowledge. Science Education, 96(6), 1071–1105. https://doi.org/10.1002/sce.21030

Osborne, J. (2014). Teaching Scientific Practices: Meeting the Challenge of Change. Journal of Science Teacher Education, 25(2), 177–196. https://doi.org/10.1007/s10972-014-9384-1

Osborne, J. F., Henderson, B. J., MacPherson, A., Szu, E., Wild, A., Yao, S.-Y., Henderson, J. B., MacPherson, A., Szu, E., Wild, A., & Yao, S.-Y. (2016). The

development and validation of a learning progression for argumentation in science. Journal of Research in Science Teaching, 53(6), 821–846.

Sampson, V., Grooms, J., & Walker, J. P. (2011). Argument-Driven Inquiry as a way to help students learn how to participate in scientific argumentation and craft written arguments: An exploratory study. Science Education, 95(2), 217–257.

Assessment of Conceptual Understanding in Physics and Conceptual Tests

Derya Kaltakçı-Gürel

Kocaeli University, Türkiye

Assessment is a process for determining the nature and extend of student learning and development. It helps us understand what students bring to class before instruction, how well they met the learning objectives after instruction, which parts of the curriculum and instruction are effective and which are not, and what unique problems our students face. In Physics Education Research (PER), research-based, research-validated conceptual assessment is crucial. Therefore, in the PER to assess conceptual understanding many different tools have been developed and used. Among them interviews, open-ended tests, and multiple-choice conceptual tests (which are usually called 'inventories') are found to be the ones commonly used in the PER to assess conceptual understanding. This review paper presented in ESERA 2023 and previously detailly discussed in the International Handbook of Physics Education Research: Teaching Physics, promises a comprehensive delve into the assessment of conceptual understanding and conceptual tests within PER.

Keywords: Assessment, conceptual understanding, conceptual tests

Assessment of Conceptual Understanding

For half a century, the investigation of students' conceptual understanding has served as a cornerstone of physics education research (PER). While a plethora of studies have delineated student comprehension of basic concepts in physics, recent inquiries have transcended the mere identification of common difficulties. Docktor and Mestre (2014) illuminate this shift, highlighting the burgeoning focus on explicating the genesis and trajectory of students' ideas - how they construct theories and explanations, and how these conceptualizations morph over time.

Despite considerable progress, a persistent chasm endures: the disconnect between pedagogical intent and actual student learning. Even seemingly elementary concepts can harbor unsuspected depths of misunderstanding. Bridging this gap necessitates precise measurement, and PER has embraced research-based, research-validated conceptual assessments as its instruments of choice.

Since the 1990s, this domain has flourished. A diversified array of validated

instruments now caters to a spectrum of topics and grade levels, propelling valuable research endeavors. From the development and validation of conceptual tests to comparative analyses of scores across demographic and cognitive metrics (e.g., culture, gender, SAT scores, reasoning ability, GPA), these assessments shed crucial light on the intricacies of the learning process.

Methods of Assessment of Conceptual Understanding

The physics education research (PER) boasts a diverse array of diagnostic tools for gauging students' conceptual understanding. These instruments - interviews, concept maps, open-ended questionnaires, word associations, drawings, essays, multiple-choice (MC) and multi-tier tests - offer nuanced insights into students' cognitive landscapes.

Within PER, interviews, open-ended questions, and MC tests reign supreme. Interviews, while yielding rich, detailed data and affording flexibility, are resource-intensive and limit generalizability due to small sample sizes. Open-ended instruments, while facilitating larger samples, incur challenges in time-consuming analysis and potential scoring ambiguities. MC tests, while efficient in administration and scoring, often fail to elicit in-depth understanding and, as Fraenkel and Wallen (2000) caution, mask the distinction between accurate reasoning and fortuitous guessing. Caleon and Subramaniam (2010b) and Peşman and Eryılmaz (2010) further highlight this limitation, emphasizing that correct answers on traditional MC tests do not necessarily equate to robust scientific conceptions, and conversely, incorrect answers may stem from valid reasoning rather than misconceptions.

To address these shortcomings, PER researchers have innovated multi-tier tests, comprising two, three, or even four tiers. These instruments leverage the strengths of various formats, mitigating the limitations of their predecessors. The multiple-tiered structure enables researchers to differentiate between correct answers gleaned from genuine understanding and those arising from chance or superficial knowledge, thus providing a more accurate picture of students' conceptualizations.

Conceptual Tests: Definition, Development Process, Interpreting the Results, Validity and Reliability Issues

Conceptual tests, also known as concept inventories, constitute research-based assessment instruments, typically employing multiple-choice formats, to diagnose students' understanding of specific physics topics. Since its seminal publication in 1992, the Force Concept Inventory (FCI) has reigned supreme as one of the most widely employed tools for gauging conceptual understanding in mechanics (Hestenes, Wells, & Swackhammer, 1992). This pioneering instrument paved

the way for a burgeoning arsenal of concept inventories specifically tailored to diverse physics domains.

Among the prominent examples enriching the PER landscape are: Test of Understanding Graphs in Kinematics (TUG-K) (Beichner, 1994), Light and Optics Conceptual Evaluation (Sokoloff, 1997), Force & Motion Conceptual Evaluation (FMCE) (Thornton & Sokoloff, 1998), Electric & Magnetic Conceptual Evaluation (EMCE) (Maloney et al., 2001), Determining and Interpreting Resistive Electric Circuits Concept Test (DIRECT) (Engelhardt & Beichner, 2004), Brief Electricity and Magnetism Assessment (BEMA) (Ding et al., 2006), Light and Spectroscopy Concept Inventory (LSCI) (Bardar et al., 2007), Mechanical Waves Conceptual Survey (Tongchai et al., 2009), and Four-Tier Geometrical Optics Test (FTGOT) (Kaltakci-Gurel et al., 2017). This non-exhaustive list exemplifies the burgeoning landscape of concept inventories embraced by PER researchers. For a more comprehensive exploration of conceptual assessment tools, PhysPort (https://www.physport.org/) and Madsen, McKagan, and Sayre (2017b) offer insightful resources. Furthermore, Kaltakci-Gurel (2023) provides a deeper exploration of conceptual tests in her 'Assessment of Conceptual Understanding" chapter within the International Handbook of Physics Education Research: Teaching Physics.

The development of a robust conceptual test necessitates adherence to a rigorous process to ensure validity and reliability. Treagust (1986) and other pioneering proponents of conceptual test development meticulously outlined a general procedure encompassing four key stages: (1) content definition, (2) student misconception identification, (3) test item development and pilot testing, and (4) test refinement. A detailed discussion of the conceptual test development process, along with the intricacies of research-based and validated instruments, is undertaken in this paper.

When conceptual tests are employed to assess conceptual understanding, the prevalent method for reporting results involves presenting frequencies or percentages within predefined conceptual categories. Two-, three-, and four-tier conceptual tests necessitate slightly nuanced approaches to calculating these percentages (Caleon & Subramaniam, 2010a; 2010b; Kaltakci-Gurel et al., 2017). Moreover, when conceptual tests are administered both before and after instruction, interpreting the outcomes and leveraging them for pedagogical improvement become paramount. Within the PER, the normalized gain (<g>), proposed by Hake (1998), has emerged as a popular, albeit approximate, metric for gauging a course's effectiveness in promoting conceptual understanding. However, in the social sciences, effect sizes, such as Cohen's d or Eta squared, often supplant the normalized gain, serving as more robust indicators of the magnitude of difference observed. Madsen, McKagan, and Sayre (2017a) elucidate the advantages of employing effect sizes. While the normalized gain

enjoys widespread use in PER, Docktor and Mestre (2014) have raised valid concerns regarding its limitations, which isa point of discussion.

Validation of conceptual tests primarily hinges on expert reviews for content validity and student interviews or factor analysis for construct validity. Liu (2012) highlights a critical issue pertaining to construct validity: the suitability of using conceptual tests for summative purposes. Unidimensionality, the extent to which a set of items measures the same underlying construct, becomes a crucial consideration for determining whether item scores can be meaningfully aggregated. The literature further reflects ongoing research interest in precisely what constructs instruments like the FCI actually measure (Hestenes & Halloun, 1995; Huffman & Heller, 1995; Lasry et al., 2011; Scott et al., 2012; Steon et al., 2020). These multifaceted issues surrounding the validity and reliability of conceptual tests is the subject of in-depth exploration.

Conclusion

Assessment has a central role in education. It offers critical insights into students' prior knowledge, mastery of learning objectives, the efficacy of curricular and pedagogical elements, and the emergence of unique student challenges. Recent research has shed light on the multifaceted nature of conceptual test outcomes, revealing their sensitivity to various facets of question design and administration. The format of assessment items (open-ended vs. multiple-choice, graphical vs. pictorial), the contextual framing (familiar vs. unfamiliar, gendered vs. neutral, self-observation vs. external observation), and even the testing conditions (timing, environment, etc.) can markedly influence student responses. These findings underscore the need for further research exploring these nuances to facilitate the development of robust and equitable conceptual assessment tools in future studies.

Recent investigations have focused on the gender gap in conceptual assessment scores, specifically the difference in performance between male and female students on standardized multiple-choice tests. Numerous studies consistently demonstrate a trend of male students outperforming female students on these assessments. This has prompted a critical inquiry: does the observed gender gap represent an artifact of the testing format itself, or does it reflect genuine differences in conceptual understanding between male and female students?

A comprehensive synthesis of research suggests that the gender gap is likely attributable to a complex interplay of factors rather than a single, isolated cause. This underscores the need for future research to deconstruct the gap and identify the specific contributing factors, particularly those related to:

- Assessment design and format: Do certain question types or test structures advantage one gender over the other?

- Prior knowledge and preparation: Do gender disparities in access to or engagement with relevant background knowledge influence performance?
- Sociocultural factors: How do societal expectations and stereotypes about gender roles in science and mathematics impact students' learning and test-taking behavior?
- Individual cognitive styles and preferences: Do males and females differ in their preferred learning approaches or problem-solving strategies, and how might these differences affect their performance on conceptual assessments?

By systematically investigating these and other potential contributing factors, future research can illuminate the true nature of the gender gap and inform the development of more equitable and effective assessment methods in physics education.

The past two decades have witnessed an exponential growth in technological advancements, leading to profound transformations in physics education, particularly in conceptual assessment and grading practices. Computers and the internet have not only reshaped students' information access but also facilitated the emergence of novel learning paradigms, such as online courses and distance learning. The limited exploration of online test administration necessitates further investigation by the physics education research (PER) community.

Physics Education Research (PER) has established itself as a vital subfield within physics, meticulously investigating diverse facets of physics education. Its focus encompasses a broad spectrum of inquiries, including the crucial domain of assessment. Within this domain, gauging students' conceptual understanding reigns supreme. This paper presented in ESERA 2023 and previously detailly discussed in the International Handbook of Physics Education Research: Teaching Physics, promises a comprehensive delve into the assessment of conceptual understanding within PER, shedding light on its current state, significant contributions, received critiques, and prospective directions.

Acknowledgement

This study mainly comes from the author's chapter called "Assessment of Conceptual Understanding in Physics" in "The International Handbook of Physics Education Research: Teaching Physics" edited by M. F. Taşar & P. R. L. Heron.

References

Bardar, E. M., Prather, E. E., Slater, T. F., & Brecher, K. (2007). Development and validation of the light and spectroscopy concept inventory. Astronomy Education Review, 5(2), 103-113.

Beichner, R. J. (1994). Testing student interpretation of kinematics graphs. American Journal of Physics, 62(8), 750-762).

Caleon, I. S., & Subramaniam, R. (2010a). Development and application of a three-tier diagnostic test to assess secondary students' understanding of waves. International Journal of Science Education, 32(7), 939-961.

Caleon, I. S., & Subramaniam, R. (2010b). Do students know what they know and what they don't know? Using a four-tier diagnostic test to assess the nature of students' alternative conceptions. Research in Science Education, 40, 313-337.

Chen, C. C., Lin, H. S., & Lin, M. L. (2002). Developing a two-tier diagnostic instrument to assess high school students' understanding- the formation of images by plane mirror. Proc. Natl. Sci. Counc. ROC(D), 12(3), 106-121.

Ding, L., Chabay, R., Sherwood, B., & Beichner, R. (2006). Evaluating an assessment tool: brief electricity and magnetism assessment. Physical Review Special Topics-Physics Education Research, 2(1), 10105-1-10105-7.

Docktor J. L. & Mestre, J. P. (2014). Synthesis of discipline-based education research in physics. Phys. Rev. ST Phys. Educ. Res. 10, 020119.

Engelhardt, P. V., & Beichner, R. J. (2004). Students' understanding of direct current resistive electric circuits. American Journal of Physics, 72(1), 98-115.

Frankel, J. R., & Wallen, N. E. (2000). How to design and evaluate research in education (4th ed.). US: McGraw-Hill Comp.

Hake, R. R. (1998). Interactive engagement vs traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. American Journal of Physics, 66, 64-74.

Hestenes, D. & Halloun, I. (1995). Interpreting the Force Concept Inventory. A response to Huffman and Heller. The Physics Teacher, 33, 502-506.

Hestenes, D., Wells, M. & Swackhamer, G. (1992). Force Concept Inventory. The Physics Teacher, 30, 141-158.

Huffman, D. & Heller, P. (1995). What does the Force Concept Inventory actually measure? The Physics Teacher, 33, 138-143.

Kaltakci-Gurel D., Eryilmaz A., McDermott L. C. (2017). Development and application of a four-tier test to assess pre-service physics teachers' misconceptions about geometrical optics. Research in Science Technological Education, 35(2), 238-260.

Kaltakci-Gurel D., Eryilmaz A., McDermott L. C. (2015). A Review and Comparison of Diagnostic Instruments to Identify Students Misconceptions in Science. EURASIA Journal of Mathematics, Science & Technology

Education, 11(5), 989-1008.

Kaltakçı Gürel, D. (2023). Assessment of Conceptual Understanding in Physics. In M. F.

Taşar & P. R. L. Heron (Eds.), The International Handbook of Physics Education Research: Teaching Physics, 19-1- 20, AIP Publishing (online), Melville, New York, available at: https:// doi.org/10.1063/9780735425712_019

Lasry, N., Rosenfield, S., Dedic, H., Dahan, A. & Reshef, O. (2011). The puzzling reliability of the Force Concept Inventory, Am. J. Phys. 79, 909.

Liu, X. (2012). Developing measurement instruments for science education research. In B. J. Fraser, K. G. Tobin & C. J. McRobbie (Eds.), Second International Handbook of Science Education (Vol2) (pp. 651-665), NY:Springer.

Madsen, S. B. McKagan, & E. C. Sayre (2017a). Best practices for administering concept inventories. The Physics Teacher, 55(9), 530-536.

Madsen, S. B. McKagan, & E. C. Sayre (2017b). Resource Letter: RBAI-1: Research-based Assessment Instruments in Physics and Astronomy, Am. J. Phys. 85,(4), 245-264.

Maloney, D. P., O'Kuma, T. L., Heiggelke, C. J., & Van Heuvelen, A. (2001). Surveying students' conceptual knowledge of electricity and magnetism. American Journal of Physics, 69(7), 12-23.

Peşman, H., & Eryılmaz, A. (2010). Development of a three-tier test to assess misconceptions about simple electric circuits. The Journal of Educational Research,103, 208-222.

PhysPort. https://www.physport.org/assessments/

Scott, T. F., Schumayer, D. & Gray, A. R. (2012)..Exploratory factor analysis of a Force Concept Inventory data set. Physical Review Special Topics-Physics Education Research,8, 020105.

Sokoloff, D. (1997). Light and Optics Conceptual Evaluation.

Steon, S.M., McDaniel, M. A., Frey, R. F., Hynes, K. M., & Cahill, M. J. (2020). Force Concept Inventory: More than just conceptual understanding. Phys. Rev. ST Phys. Educ. Res., 16, 010105.

Thornton, R. K., & Sokoloff, D. R. (1998). Assessing student learning of Newton's Laws: The force and motion conceptual evaluation and the evaluation of active learning laboratory and lecture curricula. American Journal of Physics, 66(4), 338-352.

Tongchai, A., Sharma, M. D., Johnston, I. D., Arayathanitkul, K., & Soankwan, C. (2009). Developing, evaluating and demonstrating the use of a conceptual survey in mechanical waves. International Journal of Science Education, 31(18), 2437-2457.

Treagust, D. (1986). Evaluating students' misconceptions by means of diagnostic multiple-choice items. Research in Science Education, 16, 199-207.

Formative Assessment in The Swiss Science Classroom: What Support Do Textbooks Offer?

Regula Grob[1] and Matthias von Arx[2]

[1]University of Teacher Education Lucerne, Switzerland

[2]University of Applied Sciences and Arts Northwestern Switzerland, Switzerland

The effect of formative assessment on student learning can be high (Hattie, 2009). However, it depends largely on the implementation in classroom practice: teachers have different strategies for recognising and interpreting indications of student learning and for building further teaching-learning processes on them (Ruiz-Primo, Furtak, Ayala, Yin & Shavelson, 2010). This aspect is augmented in Switzerland with the great freedom teachers have in designing lessons. Textbooks, especially at the compulsory school levels, are mainly used as a means of guiding lesson design (Gautschi, 2017). This raises the question of the extent to which science textbooks have the potential to support teachers' formative assessment practices. In addition, it must be taken into account, that support from teaching materials alone cannot guarantee the quality of formative assessment in the classroom, but that the specific implementation plays a central role (Yin et al., 2008).

Keywords: Formative assessment, textbooks

Introduction: Formative Assessment in The Science Classroom

Formative assessment is assessment for learning (Assessment Reform Group, 2002) and can be described as a cycle of four steps (Harlen, 2013): (1) The formulation and discussion of learning goals is a prerequisite for students to demonstrate their learning and for this to be diagnosed (Andrade & Valtcheva, 2009). (2) Diagnosis of students' learning can be based on different types of data: written artefacts or oral statements. The data collected must be (3) interpreted to determine where learners are in relation to the learning goals. Based on the interpretation of this data, (4) feedback is provided to the learners on how to proceed (in learner-centred activities) or how to design subsequent instruction (in teacher-centred activities; Hattie & Timperley, 2007).

Methods for formative assessment

Formative assessment methods can be classified on a spectrum from informal-flexible ("on the fly") to planned-formal (Pryor & Crossouard, 2008). This

paper focuses on planned-formal methods. These include written feedback by the teacher (Burke, 2006), peer assessment (Topping, 2003) and self-assessment (Andrade, 2010).

Formative assessment in the Swiss science classroom

An exploratory study of the teaching practice of 20 teachers from different school levels in Northwestern Switzerland showed that these teachers often (in 20 of 54 analyzed cases) did not consider all steps of the cycle described above when using formal formative methods in science subjects, but that either the assessment criteria were not clarified or there was no opportunity for the students to use the formative assessment for the further learning process (Grob et al., 2019). It was also noticeable that the competences that were formatively assessed were chosen either for practical reasons, such as the existence of rubrics, because of the easy observability of performances, or without a conscious decision. This led to some competencies being frequently assessed (e.g. conducting experiments and investigations, in 12 of 54 cases), others almost never (e.g. developing hypotheses, only 1 case; Grob et al., 2017).

Formative assessment practice internationally

International results often show similar results (Gomez & Jakobson, 2014) and are usually explained by teachers' difficulty in using formative assessment methods coherently (Gotwals et al., 2015). The identified, underlying problem is the lack of pedagocial content knowledge (Furtak et al., 2008).

Teachers are successful in implementing formal formative assessment in science, if the materials provided are accompanied by professional development (Hondrich et al., 2016). It is therefore suggested that teachers should have access to a large amount of directly applicable materials or that the transfer of existing materials to other contexts should be supported as much as possible. The finding is limited by the fact that even very specific teaching materials are implemented differently by different teachers, which could explain different effects on student learning (Yin et al., 2008).

Research question

This leads to the following guiding question: To what extent do the official science textbooks and corresponding teaching materials for the compulsory school levels of German-speaking Switzerland support the implementation of formative assessment in the classroom?

Design

In German-speaking Switzerland, there is one official science textbook for primary school level ("NaTech Primar") and two for lower secondary school

level ("NaTech S1" and "Prisma"). Apart from the textbooks themselves, they all provide teachers' commentaries and other teaching materials.

The general teachers' commentaries as well as three representative chapters per textbook and their corresponding commentaries and materials, one each with strong biology, chemistry and physics references, were selected and all passages containing the key word "formative assessment" were identified and coded.

The codes for the qualitative content analysis (Mayring, 2010) were deductively developed and specified:

- the competence(s) / concept(s) that were assessed (deductively developed from the science curriculum; D-EDK, 2016 and from Pines & West, 1986)
- the type of data on student learning that was assessed (Harlen, 2013; see theory part of this paper)
- the role of the person who assesses (Burke, 2006; Topping, 2003; Andrade, 2010; see theory part of this paper)
- the use of the feedback produced from the formative assessment (Assessment Reform Group, 2002; Hattie & Timperley, 2007; see theory part of this paper)

The two authors of the paper coded independently. One third of the material was coded by both authors, achieving an interrater agreement of 92% (Rädiker & Kuckartz, 2019) which can be considered as good.

Results

"NaTech Primar": Textbook for primary school science

The general commentaries of the textbook for primary school science mainly focuses on the role of students' preconceptions. In the three chapters analyzed (including textbook, worksheets, and chapter-specific commentaries), several ideas for formative assessment were found in each chapter. However, the focus was on self-assessment of the students' conceptual understanding towards the end of the respective chapters.

"NaTech S1": First textbook for lower secondary school science

The general commentaries of "NaTech S1" mainly focus on explaining the differences between formative and summative assessment . In the three chapters analyzed, there is always a page suggesting possible tasks, which could be used to assess the competences of the students. The main focus of the tasks lies on the competences of the curriculum. However, it is up to the teacher to decide whether the focus of the tasks should be formative or summative. Furthermore,

there are no instructions or hints on how to implement the tasks in the classroom. Secondly, in each chapter, there is a page with self-testing tasks for the students at the end of each chapter. These tasks focus on conceptual understanding.

"Prisma": Second textbook for lower secondary school science

The general teacher commentaries of the second textbook for lower secondary school level provide more extensive explanations of different learning aims formative assessment could and should focus at, on methods for formative assessment as well as on the difference between formative and summative assessment. In the textbook chapters, the teachers find a few (less than one per chapter) assessment rubrics on science-specific competences that could be directly used to formatively assess the learning progress of the students. In addition, there is a page with self-testing tasks for the students at the end of each chapter. Similarly to the other textbook for lower secondary school, these tasks mainly focus on conceptual understanding.

Interpretation and Conclusions

Diversity of examples provided

The results show that the potential for support of the formative assessment practices is only partially exploited in the textbooks analyzed: A greater diversity of assessment methods, student data as well as concepts and competences assessed (see codes specified in the design section) could and should be provided to teachers. In addition, at least some of these examples should be presented in a fully explicit version, directly applicable in the classroom, instead of a rather implicit description only. This would facilitate the implementation of meaningful formative assessment (Hondrich et al., 2016). Peer assessment, for example, although proven to be an effective method for learning in science (Scheuermann, 2017), is not exemplified in any of the textbooks.

Operationalization of competencies central to science learning

The identified examples do not represent the entire breadth of competencies and concepts that are central to science learning. Since formative assessment is highly subject-dependent (Gotwals et al., 2015; Coffey & al., 2011) and teachers' (lack of) pedagogical content knowledge can be a limiting factor (Ruiz-Primo et al., 2010; Cisterna & Gotwals, 2018), operationalizing such central competences and concepts could be of significant help in the implementation of formative assessment..

Use of data on student learning

Finally, in the textbooks analyzed, a striking lack of guidance or instruction on how formative assessment could be used to plan the next teaching and learning steps was found. This is a missed opportunity since data from teaching practice (Grob et al., 2019) showed that teacher often fail to implement this very crucial point in their science classrooms.

References

Andrade, H. (2010). Students as the definitive source of formative assessment. In H. Andrade & G. J. Cizek (Eds.), Handbook of formative assessment (S. 90–105). Routhledge.

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. Theory Into Practice, 48:1, 12–19. doi: 10.1080/00405840802577544

Assessment Reform Group (2002). Assessment for Learning: 10 principles. Research-based principles to guide classroom practice. Nuffield Foundation.

Burke, K. (2006). From standards to rubrics in 6 steps. Hawker Brownlow Education.

Cisterna, D., & Gotwals, A. W. (2018). Enactment of ongoing formative assessment: Challenges and opportunities for professional development and practice. Journal of Science Teacher Education, 29:3, 200-222. doi: 10.1080/1046560X.2018.1432227

Coffey, J.E., Hammer, D., Levin, D. M. & Grant, T. (2011). The missing disciplinary substance of formative assessment. Journal of Research in Science Teaching 48:10, 1109–1136. doi: 10.1002/tea.20440

D-EDK Deutschschweizer Erziehungsdirektoren-Konferenz (2016). Lehrplan 21. D-EDK.

Furtak, E.M., Ruiz-Primo, M.A., Shemwell, J.T., Ayala, C.C, Brandon, P., R., Shavelson, R.J. & Yin, Y (2008). On the fidelity of implementing embedded formative assessments and its relation to student learning. Applied Measurement in Education, 21:4, 360–389. doi: 10.1080/08957340802347852

Gautschi, P. (2017). Diversité ou désordre? Les moyens d'enseignement en sciences humaines et sociales en Suisse alémanique. Bulletin CIIP 2017:4, 45–51.

Gomez, M., & Jakobsson, A. (2014). Everyday classroom assessment practices in science classrooms in Sweden. Cultural Studies of Science Education, 9:4, 825–853. doi: 10.1007/s11422-014-9595-y

Gotwals, A. W., Philhower, J., Cisterna, D. & Bennett, S. (2015). Using video to examine formative assessment practices as a measure of expertise for mathematics and science teachers. [Special Issue]. International Journal of Science and Mathematics Education, 13:2, 405–423. doi: 10.1007/s10763-015-9623-8

Grob, R., Holmeier, M. & Labudde, P. (2017). Formative Assessment to Support Students' Competences in Inquiry-Based Science Education. Interdisciplinary Journal of

Problem-Based Learning. 11:2, 1–8. doi: 10.7771/1541-5015.1673

Grob, R., Holmeier, M. & Labudde, P. (2019). Analysing formal formative assessment activities in the context of inquiry at primary and upper secondary school in Switzerland. International Journal of Science Education 43:3, 407–427. doi: 10.1080/09500693.2019.1663453

Harlen, W. (2013). Assessment and inquiry-based science education: Issues in policy and practice. InterAcademy Partnership.

Hattie, J. (2009). Visible learning. A synthesis of over 800 meta-analyses relating to achievement. Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. Review of Educational Research, 77(1), 81–112. doi: 10.3102/00346543029848

Hondrich, A.L., Hertel, S., Adl-Amini, K. & Klieme, E. (2016). Implementing curriculum-embedded formative assessment in primary school science classrooms. Assessment in Education: Principles, Policy & Practice, 23:3, 353–376. doi: 10.1080/0969594X.2015.1049113

Mayring, Ph. (2010). Qualitative Inhaltsanalyse. Beltz Verlag.

Pines, L. & West, L. (1986). Conceptual understanding and science learning: An interpretation of research within a sources-of-knowledge framework. Science Education, 70:5, 583-604. https://doi.org/10.1002/sce.3730700510

Pryor, J., & Crossouard, B. (2008). A socio-cultural theorisation of formative assessment. Oxford Review of Education, 34:1, 1–20. doi: 10.1080/03054980701476386

Rädiker, S. & Kuckartz, U. (2019). Analyse qualitativer Daten mit MAXQDA. Springer

Ruiz-Primo, M. A., Furtak, E. M., Ayala, C., Yin, Y. & Shavelson, R. J. (2010). Formative assessment, motivation, and science learning. In H. Andrade & G. J. Cizek (Eds.), Handbook of formative assessment (S. 139–158). Routhledge.

Scheuermann, H. (2017). Entwicklung und Evaluation von Unterstützungsmaßnahmen zur Förderung der Variablenkontrollstrategie beim Planen von Experimenten. Logos.

Topping, K. J. (2003). Self and peer assessment in school and university: Reliability, validity and utility. In M. Segers, F. Dochy & E. Cascaller (Eds.), Optimising new modes of assessment: In search of qualities and standards (S. 55–87). Kluwer Academic Publishers.

Yin, Y., Shavelson, R.J., Ayala, C.C., Ruiz-Primo, M.A., Brandon, P.R., Furtak, E.M., Tomita, M.K., & Young, D.B. (2008). On the Impact of Formative Assessment on Student Motivation, Achievement, and Conceptual Change. Applied Measurement in Education, 21:4, 335–359. doi: 10.1080/08957340802347845

Assessment of Science Competencies in Danish Primary and Lower Secondary Education

Jørgen Løye Christiansen[1], John Andersson[2], Dorrit Hansen[1], Lars Bo Kinnerup[2] and Jørgen Haagen Petersen[1]

[1] University College Absalon, Department of Teacher Education, Denmark [2] University College Absalon, Department of Teacher Education, Denmark

This study aims to address the challenge of assessing students' science competencies in primary and lower secondary schools in Denmark. The study employs a Design Based Research (DBR) approach, where new formative assessment designs were developed and tested in collaboration with primary and lower secondary school teachers. The intervention was guided by a common understanding of the concept of competency and an understanding of what is characteristic of competence-oriented teaching. Over a period of 13 months, 28 science teachers from 18 Danish schools participated in a one-day course on the concept of competency and three workshops focusing on inquiry, modelling, and perspective-taking competencies. The teachers shared their own lesson plans and worked with teacher educators to develop and test new assessment designs. The teachers implemented the planned lesson plans and assessment designs between workshops, and teacher educators conducted observations, recorded conversations, and interviewed teachers about their perceptions and experiences of competence-based teaching and assessment. The study findings revealed that developed rubrics and forms were useful in assessing student's science competencies and improved the teachers understanding of how to design and implement formative assessment for learning.

Keywords: Science competencies, assessment.

Introduction

Although competency is seen as a meaningful goal for education in general, it was nonetheless relatively late that the concept of competency entered the Danish education system. In 2003, the so-called FNU reports presented proposals for science competencies in a Danish science education context (Andersen *et al*., 2003; Busch *et al*., 2003). In 2014, the Danish Ministry of Education introduced the *Simplified Common Goals* in an effort to make lower secondary school exams more relevant and realistic. This included a mandatory oral-practical exam for physics/chemistry, biology, and geography in 9th grade, which aimed to test students' competencies in inquiry, modelling, perspective-taking, and

communication. However, despite these efforts, current forms of assessment still appear to primarily test knowledge and skills rather than competencies (*e.g.* Dolin, 2016; Nielsen & Dolin, 2016). Therefore, assessing the science competencies of students in primary and lower secondary schools still poses challenges.

Based on the research question: *What is the characteristics of competence-oriented teaching, and which assessment tools would be beneficial in assessing students' level of competence achievement?,* we, in this study, address this challenge. In collaboration with primary and lower secondary school teachers, through studies of the concept of competency and the characteristics of competence-oriented teaching, new formative assessment (also known as assessment for learning) designs were developed and tested; designs which are expected to be more suitable in assessing students' inquiry, modelling, and perspective-taking competencies, along with their science communication skills, in the daily ongoing science education.

Method

The methodology employed in this study is based on the well-proven Design Based Research (DBR) approach, where a particular intervention is examined by continuous iteration of design, enactment, analysis, and redesign (Brown 1992). We adopted the basic assumption of DBR; that only by intervening with new designs can we develop better theories of practice at the same time as we seek to improve practice.

Over a period of 13 months, 28 science teachers from 18 Danish schools participated in first a one-day course focusing on the latest knowledge on the concept of competency, particular concerning science competencies.

As many of the participating science teachers initially struggled to view competency as something beyond knowledge + skills, we addressed this challenge, by appling Weinert's more general definition of competency to science:

> *"Science competencies are those, for individuals, available or learnable cognitive skills and abilities that, coupled with motivation, willpower, social readiness, and social skills, produce successful and responsible problem solving in variable science situations"* (Weinert, 2001, translated after Wiesner & Schreiner, 2020, added "science")

It is clear that this definition sees competency as more than the sum of knowledge and skills.

After gaining a common understanding of the concept of competency, we investigated and agreed on what characteristics competence-oriented

science teaching have, and then it became appropriate to investigate and test different formative assessment designs that are capable of assessing science competencies.

Three workshops (WS1-3) were then conducted, focusing on inquiry (WS1), modelling (WS2), and perspective-taking competencies (WS3) respectively. The communication competence was embedded in all workshops, as it was considered a fundamental aspect of the other three competencies and contributes to the students' ability on a scientific basis to argue, discuss, etc. (see Christiansen, 2022).

The participating teachers worked with teacher educators to develop new assessment designs, with the goal of assessing students' science competencies more effectively than existing practices. Between workshops, the teachers implemented the assessment designs, and teacher educators conducted observations, recorded student conversations and interviewed teachers about their perceptions of competence-based teaching, learning and assessment. With three workshops and three sessions of competence-oriented teaching and assessment, the study employed three iterations with assessment designs primarily focused on formative assessment.

Preliminary Findings

A total of 27 of the participating teachers took part in an initial interview to uncover their pre-understanding of competence-oriented science teaching and assessment of competencies. When asked, *"what do you think is characteristic of competence-oriented science teaching?"*, 6 out of 27 teachers, in different ways, expressed that it is characteristic of competence-oriented teaching that the students are the ones who are the active part (some of them adding that students must have the opportunity for independent thinking). Two of the teachers specifically mentioned that it is characteristic of competence-oriented teaching that there are real degrees of freedom/choices for the students. Three of the teachers link the concept of curiosity to competence-oriented teaching. However, overall, the majority of the participating teachers find it difficult to characterise competence-oriented science teaching.

When asked *"How do you think competence-orientated science education can be assessed?",* 9 out of 27 teachers expressed, in different ways, that it is difficult to assess competencies. For example, one teacher said, *"I actually think that's one of the things that's difficult. That's why we're here. That's what we want to get better at"*. Another teacher stated: *"that's exactly why I signed up, because I think it's really hard, I really think it's difficult"*. A third teacher has the same general opinion, albeit expressed more bluntly: *"Oh my gosh, that's why I've taken this course, isn't it? How can it be assessed? Christ, that's what's so damn difficult [.....]. I think that was a really difficult question, I don't know what to*

answer".

In five cases, the time-consuming aspect of assessments was mentioned as an obstacle to assessing science competencies.

A few teachers did relate to the relationship between current (official) assessment practices and competence-oriented assessment practices. One teacher expresses it as follows: *"I find it hard to see how you can really test students' competencies in a written test. I think the best way is to somehow collect some data through dialogue with the students"*.

Conversations as a mean of assessing competencies showed a relatively high level of agreement, with 12 out of the 27 teachers specifically mentioning conversations as a mean of assessing competencies.

However, it was clear that the majority of the participating teachers struggled to describe a competence-orientated assessment as distinct from a more traditional assessment, although some mentioned that multiple-choice assessment forms and Kahoot assessments are not the way to go.

In order to get a common understanding of what competence-oriented teaching is, we, together with The Natural Sciences Evaluation and Development Centre (NEUC), identified six criteria for competence-oriented teaching adapted by the project:

1. The teaching is based on concrete problems that engage the majority of students.
2. It is clear and explicit to all which competence is central to the teaching and how it contributes to students' overall development of that competence.
3. Collaboration and dialogue are integrated into the teaching and clearly scaffolded.
4. The teaching provides a significant degree of freedom, while catering to the different needs of students for degrees of freedom.
5. The teaching provides both time and opportunity for students to reflect and includes scaffolding processes that support student reflection.
6. The teaching clearly draws on core science knowledge and engages students in understanding how it relates to their learning.

Observations of teachers' teaching show that in several cases it has been a challenge to ensure that the implemented teaching programmes were competence-oriented according to the six criteria mentioned above. Three criteria in particular were challenging for the teachers (criteria 1, 4 and 5):

- It was often not clear to the students that the teaching was based on a concrete problem.
- The teaching did frequently not provide a sufficient degree of freedom for

the students in order to demonstrate competence.

- The teaching did frequently not provide time and opportunity for students to reflect on the process/product.

At a subsequent validation seminar, the teachers pointed out in a questionnaire that it was difficult to teach competence-orientated due to general time pressure and because they experienced a dilemma between teaching competence-orientated and ensuring students' knowledge and skills through more traditional teaching. At the seminar, teachers confirmed that two of the criteria for competence-orientated science teaching were particularly difficult to work with. These were the criteria relating to students' degrees of freedom and time to think. These were also two of the three criteria that were rated particularly low in the observations.

Formative assessment

Through the three DBR iterations assessment designs were developed and tested. As teachers may not always be able to observe the work of all student groups, we found that continuous self (and/or peer)-assessments in student groups, were an effective formative assessment tool for assessing competencies. Such an approach enabled both teachers and students to track progress in the science competence in question as the work progressed.

In the two first examples below, students receive a document with questions to be answered. In the first one, representing aspects of the inquiry competence, students provides answers to the different inquiry questions as they work through their inquiry design in their groups. In this example (figure 1), if an answer has been written, it is marked with a colour (blue) (Of course, originally there will be written letters in the box). If the students improve/change their answer in a subsequent lesson, it is here marked with a darker shade of blue. This way, older answers are not deleted, but newer ones are entered under the lesson where the answer was completed. In this example, the group answers all the survey questions after lesson four. Both students and teachers can then visually see which answers have been changed along the way and discuss why this happened and discuss the quality of the different answers and consider improvements should a similar challenge arise in a future teaching situation.

Figure 1. Student Document – Create your own science investigation/ inquiry.

Create your own science investigation - a document for students

	First lesson	Second lesson	Third lesson	Fourth lesson
1. The science question to be investigated is as follows:				
2. Why do you want to investigate it?				
3. How and in what order will you get the information you need to answer your science question?				
4. How will you keep track of your information, i.e. record your collected data?				
5. How will you analyse your data?				
6. What was the result (conclusion) of your study?				
7. How can you argue in favour of your conclusion based on the data collected?				
8. How will you present your results?				
9. If you had the opportunity to improve your study, what would you do?				

Some teachers conducted educational situations that were basically based on an assessment structure as presented in Figure 1, but allowed students to develop their own criteria for what constitutes a good science inquiry. First, the students discussed among themselves what characterises a scientific inquiry in general and later more specifically a fieldwork inquiry. Then, in dialogue with the teacher, the students drew up a joint list of criteria for a good field inquiry. The criteria became more relevant for the students because they formulated them in their own language. The criteria were used as a guideline for the students, both when they were in the process of planning their study and along the way, when they carried out the inquiry and finally when they had to assess their results. The fact that the students themselves set the criteria against which their inquiry was measured created a motivating learning environment.

The second example (figure 2), taken from the work with the perspective-taking competence, is comparable to the previous example. Here, however, the students are asked to write in the same box, but with different font colours in the different lessons. However, the dialogue between students and teacher and between students will be comparable to the example above.

Figure 2. Student Document – Perspective-taking

Perspective-taking – Student document

A little about putting things into perspective:

Perspective-taking is about putting something into a larger perspective, going beyond your study/results.

Perspective-taking is about putting your topic/results into other contexts, i.e. relating it to something else.

Fill out this form in your work with the perspective-taking competency:

Things you need to consider:	Your answers:	
Which study or results are the starting point for your perspective:	XXXXXXXXXXXXXXXXXXXXXX (LESSON ONE)	
How can you put that into perspective:	Your answer:	Reason for the answer (it is because....)
• in your everyday life? (Neighbourhood?):	XXXXXXXXXXXXXXXXXXXX (LESSON ONE) XXXXXXXXXXXXXXXXXXXX (LESSON TWO) XXXXXXXXXXXXXXXXXXXX (LESSON THREE)	XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXX
• in relation to other regions:		
• in relation to other science issues:		
• in relation to social issues:		
• in relation to conflicts of interest		
• Time-wise (the same 'subject area' but in a different time; past or future):		

Forms with relevant questions that ensured that students argued on a scientific basis was also tested in the activity "design a model". Here students would have to consider; what is the aim of your model? (Answer and justify your answer), what is the model for? What should it not be used for? Who is the target audience of the model? To what extent does the model reach the target audience? etc. Based on the above questions, one of the schools developed a self-assessment form for lower secondary education (see figure 3). Here, the students could evidently see the teachings assessment objectives and could tick the box on the form where they thought they were placed at a certain point in time. The tick marks were then the starting point for dialogue within the group and with the teacher. Was it possible to argue in favour of the position? And what kind of work and insight was needed to be able to tick higher up the hierarchy, so to speak?

Figure 3. Student Document – Self-assessment Designing Models.

Modelling: Designing models				
Grade level	**Adequate (02-4)**	**Fairly good (4-7)**	**Very good (7-10)**	**Excellent (10-12)**
Purpose	Can describe the purpose of own model with help	Can describe the purpose of own model with increasing complexity		Can explain the purpose of own model
Construction	Can reproduce a model with simple modifications	Can reproduce a model with increased level of detail	Can combine elements from different models into their own model	Can independently plan and design own model
Usage	Can identify elements from own model using simple relevant terminology	Can describe the coherence of own model using simple relevant terminology	Can explain correlations and details in their own model using relevant terminology	Can use own model to predict (put into perspective) specific issues using relevant terminology
Valuation	Can identify simple differences between own model and reality	Can compare own model with similar models	Can mention advantages and disadvantages of own model	Can make reasoned suggestions for improving own model

Discussion and Conclusion

Formative assessment enriches the process of learning by promoting dynamic dialogues and collaborative interactions between students and educators, thereby advancing the learning process (see Arnold, 2022). Teachers must observe how their students react to different forms of feedback to customize it to meet each student's specific needs. However, it is challenging the teacher to gain insight into students' way of thinking and to frame feedback that guides them towards specific learning goals (Black & Wiliam, 2009). Student self-assessment becomes a potent tool for enhancing student achievement, especially when students are trained in utilizing predefined sets of performance criteria like rubrics to assess their own work, or when they receive other direct instruction on self-assessment (Ross, 2006).

In the context of peer assessment, students take on roles akin to instructional resources for one another, similar to collaborative learning. The feedback exchanged among peers during this assessment process provides valuable insights into their comprehension levels (Black & Wiliam, 2009).

For students to embrace these roles effectively, it is crucial that they possess a clear understanding of both the learning objectives and the performance criteria. Some strategies to facilitate this include teachers and students collaboratively constructing assessment criteria to enhance the reliability of peer assessment (Topping, 2010), or teachers demonstrating the process to students to encourage

their active participation (Black & Wiliam, 2009).

Assessment of competency in education requires, as we see it, a framework of unknown and uncertainty (cf. definition of competency). Therefore, teaching must be organized in a way that allows for a significant degree of freedom for students, with opportunities for collaboration and dialogue integrated into the teaching and clearly scaffolded by the teacher. This provides the basis for assessing competencies, when learners is observed in specific and authentic situations. In this project, this was typically conducted within the context of problem-based learning and group work. Students formulated, for example, their hypotheses or research questions, and designed and explored their own self-chosen path from question to answer. Through this dialogue-based approach, teachers scaffolded students by asking questions rather than providing answers, promoting scientific thinking and argumentation.

However, the teacher may not always be able to fully observe the work of all student groups. To address this, this study found that continuous self-assessments in student groups, largely based on rubrics and forms, were an effective formative assessment tool for assessing competencies. Such an approach enabled both teachers and students to track progress in the competency as the work progressed.

As mentioned, competence-oriented science teaching was a challenge for many of the participating teachers in the project, even though they are assumed to have special academic qualifications and consider competence-oriented teaching important. The analysis found some of the six criteria that were rated low, indicating that the teaching was not sufficiently competence-oriented.

This also suggests that there is still a need for efforts that can support more teachers to be able to work in a competence-oriented way and that assessment practices support competence-oriented science teaching.

Acknowledgements

This work was carried out with financial support of the Novo Nordisk Foundation grant no. NNF20OC0065745.

References

Andersen, N.O., Busch, H., Troelsen, R. & Horst, S. (2003). Fremtidens naturfaglige uddannelser [Future science education]. Uddannelsesstyrelsens temahæfteserie, nr. 7, 2003. København: Undervisningsministeriet.

Arnold, J. (2022). Prioritising students in Assessment for Learning: A scoping review of research on students' classroom experience. Review of Education, 10, e3366. 36 pp. https://doi.org/10.1002/rev3.3366.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment.

Educational Assessment, Evaluation, and Accountability, 21(1), 5–31. https://doi.org/10.1007/s11092-008-9068-5

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. Journal of the Learning Sciences, 2(22), 141-178. https://doi.org/10.1207/s15327809jls0202_2

Busch, H., Horst, S. & Troelsen, R. (2003) (eds.). Inspiration til fremtidens naturfaglige uddannelser [Inspiration for future science education]. København: Undervisningsministeriet.

Christiansen, J.L. (2022): Naturfaglige kompetencer i grundskolen – hvilke bør der fokuseres på? [Science competencies in primary and lower secondary schools - which ones should be prioritised?] MONA, 2022 (3): 6-17.

Dolin, J. (2014): Naturfaglige kompetencer - om kompetencetænkningen i nye Forenklede Fælles Mål [Competencies in the science education - on the concept of competencies in the new Simplified Common Goals]. *In* Tougaard,S. & Kofod, L.H. (eds.): Metoder i naturfag – en antologi. Experimentarium, København; 49-66.

Dolin, J. (2016): Idealer og realiteter i målorienteret undervisning [Ideals and realities in aim-oriented education]. *In* Krogh, E. & Holgersen, S.E. (2016) (eds.). Sammenlignende fagdidaktik 34. Cursiv 19. pp. 67-87.

Nielsen, J.A. & Dolin, J. (2016). Evaluering mellem mestring og præstation [Assessment between mastery and performance]. MONA, 2016 (1): 51-62.

Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. Practical Assessment, Research, and Evaluation: Vol. 11, Article 10. https://doi.org/10.7275/9wph-vv65

Topping, K. J. (2010). Peers as a source of formative assessment. In H. L. Andrade and G J. Cizek (Eds.), Handbook of formative assessment (pp. 61–74). New York: Routledge. https://doi.org/10.5937/IJCRSEE2002121C

Wiesner, C. & Schreiner, C. (2020). Ein Modell für den kompetenzorientierten Unterricht und als Impuls für reflexive Unterrichtsentwicklung und –forschung [A model for competence-oriented teaching and as an impulse for reflective teaching development and research]. *In* Greiner, U., Hofmann, F., Schreiner, C. & Wiesner, C. (eds.), Bildungsstandards: Kompetenzorientierung, Aufgabenkultur und Qualitätsentwicklung im Schulsystem (pp. 319-352). Waxmann. https://doi.org/10.31244/9783830991489

Investigating Gender Differences in Confidence and Accuracy in the Force Concept Inventory (FCI)

Derya Kaltakçı-Gürel

Kocaeli University, Türkiye

Confidence is defined as students' belief in their own ability. According to the confidence paradigm students report how confident they are in the accuracy of their responses in an exam. It is considered as mis-calibration if there is a mismatch between accuracy and confidence. Bias is obtained by subtracting the mean accuracy score from the mean confidence score. Previous research revealed that individual differences such as gender, influences confidence and bias scores. The aim of the present study is to investigate the effect of gender on freshman students' performance and confidence in the Force Concept Inventory (FCI). Eighteen items from the FCI were administered to 369 (n_{female}=185, n_{male}=184) freshman physics students in a state university in Türkiye. Each FCI item was featured with an item asking students' confidence in answering the item correctly. The independent samples t-test was conducted to compare the confidence scores and FCI scores for male and female students separately. Results indicate that gender differences were exist both on confidence and FCI scores. Whereas the effect sizes were found to be medium in confidence and small in the FCI scores, respectively. The implications of the results for teaching and learning is discussed in the study.

Keywords: Bias; confidence; FCI; gender

Introduction

Since the 1970s, individuals have been asked to rate their confidence in the accuracy of their responses to test items as one of the primary methods of examination in calibration research in the discipline of psychology (Echternacht, 1972; Stankov & Crawford, 1997). Confidence ratings are useful for reducing the impacts of guessing and enhancing the test's reliability through its unique scoring method. In the confidence testing scoring methods, a correct response provided with high confidence received more credit than a false response provided with low confidence (Echternacht, 1972).

Confidence is an essential metacognitive construct that concerns the self-assessment of one's own knowledge. When a student feels more confident in his/her responses when they are correct than when they are incorrect, it is seen as having a good calibration in the confidence rating; however, it is regarded as having been mis-calibrated if there was a mismatch between correctness and confidence (Baker & Fogarty, 2004; Baker, 2010; Lundeberg, 1992; Lundeberg,

Brown & Elbedour, 2000; Caleon & Subramaniam, 2010b). Bias score indicated the size and direction of this mismatch and was calculated as the average of confidence ratings across all items in a task minus the proportion correct score for that task. (Baker & Fogarty, 2004; Stankov & Crawford, 1997). Overconfidence exists if a bias score has a positive sign, underconfidence does if it has a negative sign, and perfect calibration is shown when the bias score is zero. Confidence also used in item analysis during test development. Poor calibration of an item may indicate problems in wording of the item or may point to the presence of an attractive alternative among choices (Stankov, Lee, Luo & Hogan, 2012).

According to research in this area, individual differences in academic performance, gender, age, personality, or ability can affect confidence and bias ratings (Baker, 2010; Lundeberg, 1992; Stankov & Crawford, 1997). As an example, females were found to have lower levels of confidence in their mathematical and scientific abilities than males, and this gender disparity worsens with age (Hyde, Fennema, & Lamon, 1990; Lundeberg & Moch, 1985 as cited in Lundeberg, Fox & Puncohar, 1994; Lundeberg et al., 2000).

Al-Rubeyea (1996), Caleon and Subramaniam (2010), Cataloglu (2002), Franklin, (1992), Hasan, Bagayoko and Kelley (1999), Kaltakci Gurel, Eryilmaz and McDermott, (2017) and Woehl (2010) made a commendable effort in using confidence in their test development. Most of these studies used confidence ratings and conceptual performance on specific topics to explore the strength of misconceptions in those specific physics topics. Although the confidence paradigm has been extensively studied in psychology, there is little in science education (Caleon & Subramaniam, 2010; Testa, Colantonio, Galano, Marzoli, Trani and di Uccio, 2020). Sharma and Bewes (2011), Jack, Liu, Chiu and Tsai (2012), Leppavirta (2012), Rachmatullah and Ha (2019), Testa et al. (2020) reported valuable efforts about the relationships between confidence and understanding as well as the effects of culture, gender and instruction. The present study tries to investigate the gender differences, if any, in confidence and accuracy in a well-known conceptual test FCI. The research questions that addressed in the present study were:

- Is there a significant difference in the mean confidence scores for male and female students?
- Is there a significant difference in the mean FCI scores for male and female students?

Methods

Participants

Data of this study were obtained from 369 first year undergraduate physics

classes at a state university in Türkiye. The sample consisted of 185 (50.1%) male and 184 (49.9%) female students as shown in Table 1.

Table 1. Sample of the study.

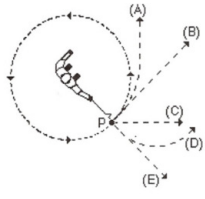
Gender	Frequency (f)	Percentage (%)
Male	185	50.1
Female	184	49.9
Total	369	100

Data Collection Tool

Force Concept Inventory (FCI) was developed by Hestenes, Wells and Swackhammer (1992) to assess conceptual understanding in physics. The original FCI contains 30 multiple choice items in mechanics and kinematics. In the present study 18 items from the FCI was used as a measurement tool. A Turkish translation of the FCI by Temizkan (2003) was used. Each FCI item was featured with an item asking students' confidence in answering the item correctly.

Figure 1. A Sample Item from FCI and Confidence Item

A steel ball is attached to a string and is swung in a circular path in a horizontal plane as illustrated in the accompanying figure.
At the point P indicated in the figure, the string suddenly breaks near the ball.
If these events are observed from directly above as in the figure, which path would the ball most closely follow after the string breaks?



How confident are you with your answer in the above question?

A. Strongly not confident
B. Not confident
C. Fairly confident
D. Confident
E. Strongly confident

Data Analysis

The independent samples t-tests were used to answer the research questions to uncover the differences between male and female students in confidence and FCI scores. IBM SPSS Statistics version 23 was used to perform all analyses.

Results

An independent samples t-test was calculated to compare the confidence scores for male and female students. According to the statistical analysis, there was a significant difference in confidence scores for females (M=53.51, SD=28.58) and males (M=66.99, SD=26.67; t=-4,68, p=0.00). The magnitude of the difference in the confidence score means was moderate (eta squared=0.06).

Based on the independent samples t-test to compare the FCI scores for male and female students, there was a significant difference in FCI scores for females (M=49.69, SD=22.52) and males (M=56.98, SD=24.21; t=-2.99, p=0.003). The magnitude of the difference in the FCI means was small (eta squared=0.02).

Discussion and Conclusion

When the confidence does not correspond to the accuracy, overestimation occurs. This is called as overconfidence bias. The male freshman physics students' confidence and FCI accuracy scores were found significantly higher than the ones for the females. The effect size calculations showed that the difference was more for confidence scores than the FCI scores. These results are similar to the previous studies (Kaltakci-Gurel, 2023; Sharma & Bewes, 2011), reporting males were found to be more successful and confident in their responses than females in physics. Conceptual understanding is thought to be related to both correctly answering a question and recognizing the correctness of the answer at a metacognitive level. One's ability to assess own knowledge and decisions is considered an important skill in educational endeavors (Kaltakci-Gurel, 2023). As a result of instruction, students are expected to improve their ability to assess their performance, besides moving from naive conceptions to the scientific models. However, even after a great deal of courses taken, students still have deficiencies in their conceptual understanding. Programs at the conceptual level is essential to teach physics, but not sufficient. Female students' confidence in their conceptual knowledge should be increased. The results discussed here are limited to the participants and the context within which this research was conducted. For the future research, the next step might be a more in-depth investigation of why gender differences exist in conceptual understanding and confidence in qualitative manner.

References

Al-Rubayea, A. M. (1996). *An Analysis of Saudi Arabian High School Students,, Misconceptions About Physics Concepts*. Unpublished doctoral dissertation, Kansas State University, Manhattan, Kansas.

Baker, S. F. & Fogarty, G. J. (2004). *Confidence in cognition and intrapersonal perception: Do we know what we think we know about our own cognitive performance and personality traits?*. In M. Katsikitis (Ed.), Proceedings of the 39th Australian Psychological Society Annual Conference: Psychological Science in Action, Australian Psychology Society (pp.24-28), Melbourne, Australia.

Baker, S. F. (2010). *Calibration analysis within the cognitive and personality domains: Individual differences in confidence, accuracy, and bias.* Unpublished PhD Thesis, University of Southern Queensland, Retrieved April 23, 2012, from http://eprints.usq.edu.au/8856/2/Baker_2010_whole.pdf

Caleon, I. S., & Subramaniam, R. (2010b). Do students know what they know and what

they don't know? Using a four-tier diagnostic test to assess the nature of students,, alternative conceptions. *Research in Science Education, 40*, 313-337.

Çataloğlu, E. (2002*). Development and validation of an achievement test in introductory quantum mechanics: The quantum mechanics visualization instrument (QMVI).* Unpublished doctoral dissertation, The Pennsylvania State University, Pennsylvania.

Echternacht, G. J. (1972). The use of confidence testing in objective tests. *Review of Educational Research, 42*(2), 217-236.

Franklin, B. J. (1992). *The development, validation, and application of a two-tier diagnostic instrument to detect misconceptions in the areas of force, heat, light and electricity.* Unpublished PhD Thesis, The Louisiana State University.

Hasan, S., Bagayoko, D., & Kelley, E. L. (1999). Misconceptions and the certainty of response index (CRI). *Physics Education, 34*(5), 294-299.

Hestenes, D., Wells, M. & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher, 30*, 141-158.

Kaltakci-Gurel, D., Eryılmaz, E. & McDermott, L. C. (2017). Development and application of a four-tier to assess pre-service physics teachers' misconceptions about geometrical optics. *Research in Science & Technological Education, 35*(2), 238-260.

Kaltakci-Gurel, D. (2023). Exploring pre-service teachers' conceptual understanding and confidence in geometrical optics: A focus on gender and prior course achievement. *Education Sciences, 13*, 452, 1-11. https://doi.org/10.3390/educsci13050452

Leppavirta, J. (2012). Assessing undergraduate students' conceptual understanding and confidence of electromagnetics. *International Journal of Science and Mathematics Education, 10,* 1099–1117.

Lundeberg, M. A., Fox, P. W., & Puncochar, J. (1994). Highly confident but wrong: gender differences and similarities in confidence judgments. *Journal of Educational Psychology, 86*(1), 114-121.

Lundeberg, M. A., Fox, P. V., Brown, A. C., & Elbedour, S. (2000). Cultural influences on confidence: country and gender. *Journal of Educational Psychology, 92*(1), 152-159.

Rachmatullah, A. & Ha, M. (2019). Examining high-school students' overconfidence bias in biology exam: a focus on the effects of country and gender. *International Journal of Science Education, 41*(5), 652-673.

Sharma, M. D. & Bewes, J. (2011). Self-monitoring: Confidence, academic achievement and gender differences in physics. *Journal of Learning Design, 4*(3), 1-13.

Stankov, L., & Crawford, J. D. (1997). Self-confidence and performance on tests of cognitive abilities. *Intelligence, 25*(2), 93-109.

Stankov, L., Lee, J., Luo, W. & Hogan, H.J. (2012). Confidence: A better predictor of academic achievement than self-efficacy, self-concept and anxiety? *Learning Individual Differences 22* (6), 747–758.

Temizkan, D. (2003). *The effect of gender on different categories of students'*

misconceptions about force and motion. Unpublished master thesis, Middle East Technical University, Ankara, Türkiye.

Testa, I, Colantonio, A., Galano, S., Marzoli, I., Trani, F. & di Uccio, U. S. (2020). Effects of instruction on students' overconfidence in introductory quantum mechanics. *Phys. Rev. Phys. Educ. Res. 16*, 010143.

Woehl, K. A. F. (2010). *The effects of practice questions and accuracy assessment experience on metacognition.* Unpublished PhD Thesis, University of North Dakota: Grand Forks.

Characteristics of Written Examination Questions That Differentiate Between Students Who Have Completed Practical Work in Hands-On and Other Ways in Science Lessons

Alistair Moore[1], Peter Fairhurst[1], Judith Bennett[1], Christine Harrison[2], Catarina Correia[3], Jessie Durk[4]

[1] University of York, United Kingdom

[2] King's College London, United Kingdom

[3] University College London, United Kingdom

[4] Imperial College London, United Kingdom

Teachers' decisions about the use of laboratory and other practical learning activities in science lessons are influenced by national, high-stakes, summative assessments of the skills and understanding developed through these activities. It is important that such assessments are constructed to reward and incentivize effective practices in practical work in lessons. In England, 16-year-olds' knowledge and understanding of practical work in science are now summatively assessed only via written questions in high-stakes national examinations. This quantitative, empirical study investigated whether sets of these written questions with different characteristics could differentiate between students who had experienced practical activities only through either hands-on work, teacher demonstration, video demonstration, or reading about the activity. Conclusions are drawn from 1486 post-intervention tests, comprising written questions drawn from national assessments, completed by students aged 14-15 in England. We identify characteristics of written questions that did and did not differentiate between students in the four practical intervention groups, to establish a basis for recommendations to refine such questions in national examinations to better reward and incentivize effective practices (i.e. practices that better support learning) in practical work in science lessons.

Keywords: Assessment; practical work; washback

Introduction

School science lessons often incorporate laboratory and other practical work (activities that combine interaction with materials and apparatus with aspects of experimenting, observing, measuring, investigating and analysing data). Students may engage with a practical activity through hands-on participation,

but they may also accumulate varying amounts of assessable knowledge and understanding of it by watching a teacher demonstration of it, or watching a video of it, or reading a written account of it (Moore et al., 2020).

The purposes of practical work and its effectiveness in supporting learning have been debated (e.g Abrahams & Millar, 2008; Gericke et al., 2023), as has the validity of assessments of the skills and understanding developed through practical work (e.g. Black et al., 2010; Childs & Baird, 2020). Abrahams et al. (2013) differentiate between *direct* assessment, in which students' practical skills are assessed during hands-on practical work, and *indirect* assessment, in which students' practical skills are inferred from data they collected or from their written account of practical work. Since 2018 in England, the skills, knowledge and understanding that 16-year-olds have gained through practical work are now summatively assessed only indirectly, via written questions in high-stakes national examinations.

Teachers' decisions about the inclusion of practical activities in lessons, and whether these are hands-on, are influenced by numerous factors, including the nature of associated assessments. Assessments, particularly high-stakes summative assessments, influence teaching through 'washback' effects (Cheng et al., 2015). Science teachers' choices in their use of practical activities is 'routinely influenced' by summative assessments (Abrahams & Saglam, 2010). High-stakes assessments at age 16 in England have 'narrowed the ways practical work is conducted' (Childs & Baird, 2020), and increasing reliance upon written assessments 'reduces the likelihood that practical work will be taught and learnt as well as it might be' (Abrahams et al., 2013). Therefore, where written assessments are the main or only summative assessment of practical skills and understanding, it is crucial that their construction helps to incentivise – by rewarding – effective practices in practical work in lessons; to do that the assessments must differentiate between students who have experienced practical work in different ways.

This study investigated the ability of written examination questions from high-stakes, summative assessments in England to differentiate between students who had experienced practical activities in hands-on and other ways (research question 1).

Theoretical Framework

Purposes of Practical Work

Broadly, practical work has been said to have the following purposes: (1) to develop understanding of phenomena and scientific explanations; (2) to develop understanding of scientific approaches to enquiry; (3) to develop competencies in the use of apparatus and procedures; (4) to develop transferable skills such

as objectivity and teamwork; (5) to increase motivation and engagement with science (Holman, 2017; Millar & Abrahams, 2009; Wellcome Trust, 2020). Practical work may be deemed 'effective' if it helps students to achieve learning outcomes related to at least one of these purposes.

Assessment of Practical Skills

With reference to the three domains of Bloom's taxonomy, the above purposes of practical work span the cognitive, psychomotor and affective domains, but only the first purpose is wholly within the cognitive domain. Learning outcomes in the psychomotor domain can only be assessed directly while pupils perform practical tasks (Abrahams et al., 2013), and outcomes in the affective domain (such as increased motivation) can only be inferred from other behaviours (Gauld & Hukins, 1980). Previous analysis has suggested that written examination questions set in practical contexts only assess outcomes in the cognitive domain (Bennett & Kennedy, 2001); hence, they are likely to only assess outcomes related to purpose 1, and parts of purposes 2 and 3, above. Thus, practical work may be deemed 'effective' as a preparation for examination questions if it helps students to answer questions assessing learning outcomes related to these purposes.

Different orders of thinking skills within Bloom's cognitive domain include recall, application and analysis. The assessment organisations that set high-stakes examinations in England categorise questions according to which of these orders of thinking skills they assess. Specifically, questions assessing knowledge and understanding of practical work may assess recall of learned details of practical techniques and procedures, application of practical knowledge and understanding in unfamiliar scenarios, or analysis of presented information, including interpretation, evaluation and drawing conclusions. This categorisation of questions was adopted as a framework for analysis in this study so that the findings could be communicated to the awarding organisations in their own language, to maximise potential impact.

This study investigated whether there were differences in the abilities of sets of written examination questions assessing recall, application and analysis to differentiate between students who had experienced practical activities in hands-on and other ways (research question 2).

Research Questions

1. Can written examination questions differentiate between (by differentially rewarding) students who have completed practical activities in hands-on and other ways?
2. What are the generalizable characteristics of written questions that differentiate between these students?

Methods

In this empirical, quantitative study, data were collected for six practical activities common in school science courses at age 14-16 in England. These were: measuring the distribution and abundance of plants in a field using quadrats; determining the rate of an enzymatic reaction (catalase and hydrogen peroxide) by measuring the volume of gas produced over time; reacting to excess, separation and crystallization while making copper sulfate salt; determining the rate of a chemical reaction (sodium thiosulfate and hydrochloric acid) be measuring the time taken to reach an end point ('disappearing cross' method); measuring the acceleration of a trolley down a slope; and measuring the extension of a suspended spring caused by attaching masses to it.

1911 students aged 14-15 in two regions of England participated in the study over 3 consecutive school years. Four interventions were compared for each practical activity: students either participated in a hands-on practical version of the activity, or watched a teacher demonstration of it, or watched a video demonstration of it, or read a written description of it. Cohorts of students in the intervention groups were balanced on characteristics including predicted examination grades, gender, ethnicity, socioeconomic factors and the teacher's level of experience.

Quantitative data on pupil performance were collected from 1486 post-intervention tests comprising written questions assessing knowledge and understanding of the practical activity, derived from national, high-stakes summative examinations in England. Students in all four intervention groups sat the same post-intervention test for each practical activity.

To investigate characteristics of written questions that did and did not differentiate between students in the intervention groups, subsets of questions assessing different orders of thinking skills in Bloom's cognitive domain were analyzed. Subsets of questions requiring students to answer via multiple-choice, short written answer, extended written answer, and calculation were also analyzed.

Quantitative data analysis aimed to investigate whether the test scores on sets of questions with particular characteristics was affected by the intervention type. When comparing the four different interventions for the same practical activity, test marks were compared using a classic ANOVA test and a Tukey HSE post-hoc test for homogenous data sets, or for non-homogenous data sets a Welch ANOVA test and Games-Howell post-hoc test. A Kruskal-Wallis test for non-parametric data was used to confirm any statistically significant differences in mean test scores.

Results

At whole-test level, the only statistically significant difference ($p=0.05$) in mean post-intervention test score was between the teacher demonstration intervention group (52.2% ± 3.8%; mean ± 95% confidence interval) and the video group (43.6% ± 3.5%), with a large effect size (0.99; Cohen's d). Differences between every other pair of interventions were not significant at whole-test level.

The post-intervention tests comprised a mix of written questions assessing different orders of thinking skills within Bloom's cognitive domain, and requiring students to answer in different ways. Analysis of subsets of questions with these characteristics indicated that questions with particular characteristics were more able than others to differentiate between students in the four different intervention groups, as follows.

In general, questions assessing the thinking skills of application or analysis were more likely to differentiate between the intervention groups than questions assessing only recall. This pattern was observed for subsets of questions requiring students to answer in different ways, including questions requiring an extended written answer and multiple-choice questions. For questions requiring an extended written answer (worth 6 marks), there were no statistically significant differences between the mean post-intervention test scores for the four intervention groups. When questions requiring an extended written answer assessed a mixture of recall, application and analysis, there was a statistically significant difference between some of the intervention groups as determined by one-way Welch ANOVA ($F(3,233) = 6.591$, $p < 0.001$). Multiple-choice questions did not differentiate between intervention groups when only recall was assessed, as determined by one-way ANOVA ($F(3,477) = 1.707$, $p = 0.165$); but they did differentiate between intervention groups when application or analysis was assessed, as determined by one-way Welch ANOVA ($F(3,448) = 10.303$, $p = <0.001$).

The picture was more nuanced for questions requiring a short written answer (worth 1-3 marks). These questions did not differentiate between the intervention groups when they assessed only recall of *what* was done (e.g. recall of steps in a practical procedure), but did differentiate when they assessed recall of *why* particular things were done (e.g. to increase safety or accuracy). When these questions assessed application or analysis they did differentiate between the intervention groups, except when the answers could be deduced from a provided diagram with little or no understanding of the practical activity.

Questions assessing mathematical skills did, in general, differentiate between the intervention groups, except when the question had what McAlinden & Noyes (2019) call low 'entanglement' with the practical context (i.e. the question could be answered using rehearsed mathematical skills alone with little or no

understanding of the practical activity).

Conclusions and Implications

This study has identified generalizable characteristics of written examination questions that are more likely to differentiate between (by differentially rewarding) students with different experiences of practical work.

In general, questions assessing only recall of learned details of practical activities were less likely to differentiate between students who had experienced the practical activities in hands-on and other ways. The findings suggest that having experienced the practical activity through hand-on participation or by watching a high-quality teacher demonstration of it conferred no significant advantage in answering questions assessing only recall, and having experienced the practical activity only by watching a video of it or reading about it conferred no significant disadvantage.

In general, questions assessing application of practical knowledge and understanding in unfamiliar scenarios, or analysis of presented information (including interpretation, evaluation and drawing conclusions), were more likely to differentiate between students who had experienced practical activities in hands-on and other ways. In these cases, having experienced the practical activity through hand-on participation or by watching a high-quality teacher demonstration of it conferred an advantage in answering the questions.

The format in which students were required to give their answer, be it via multiple choice, short or extended written answer, or calculation, also affected the ability of the question to differentiate between students with the different practical experiences.

These findings are important as they suggest ways in which high-stakes, summative examination papers could be optimized by including practical questions with combinations of the characteristics we have identified, to better reward students who have had particular practical experiences – and therefore, to better incentivize these kinds of practical experiences in lessons. This could create positive washback effects on teaching and learning. The findings have been shared with the assessment organisations in England.

Further Work

Additional, qualitative data collected by us from lesson observations and teacher interviews suggests that questions assessing application and analysis skills better rewarded practical experiences with pedagogical features that were more noticeable in teacher demonstrations and hands-on practical work. These features included increased presence of teacher guidance (especially teacher-student dialogue and focused questioning) and actively-engaged students. These

findings are discussed in detail elsewhere (Moore et al., 2023).

We suggest that practical experiences with these features better support learning and development of skills related to all the aforementioned purposes of practical work. Therefore, even though written assessments can only assess learning outcomes related to a subset of these purposes, it may nevertheless be possible to optimize the assessments to better reward, and therefore better incentivize, the kinds of practical experiences that support learning related to all the purposes of practical work.

Acknowledgements

This study was funded by the Wellcome Trust, the Gatsby Charitable Foundation and the Royal Society (Wellcome Trust grant number 208491/Z/17/Z). The authors acknowledge the contributions of Lucy Wood and Nicklas Lindstrom (King's College London) during data collection, thank Ginny Page (Gatsby) and Mat Hickman (Wellcome) for their support, and thank all the teachers and students who participated in this study.

References

Abrahams, I., & Millar, R. (2008). Does practical work really work? A study of the effectiveness of practical work as a teaching and learning method in school science. *International Journal of Science Education*, *30*(14), 1945-1969.

Abrahams, I., Reiss, M. J., & Sharpe, R. M. (2013). The assessment of practical work in school science. *Studies in Science Education*, *49*(2), 209-251.

Abrahams, I., & Saglam, M. (2010). A study of teachers' views on practical work in secondary schools in England and Wales. *International Journal of Science Education*, *32*(6), 753-768.

Bennett, J., & Kennedy, D. (2001). Practical work at the upper high school level: the evaluation of a new model of assessment. *International Journal of Science Education*, *23*(1), 97-110.

Black, P., Harrison, C., Hodgen, J., Marshall, B., & Serret, N. (2010). Validity in teachers' summative assessments. *Assessment in Education: Principles, Policy & Practice*, *17*(2), 215-232.

Cheng, L., Sun, Y., & Ma, J. (2015). Review of washback research literature within Kane's argument-based validation framework. *Language Teaching*, *48*(4), 436-470.

Childs, A., & Baird, J. A. (2020). General Certificate of Secondary Education (GCSE) and the assessment of science practical work: an historical review of assessment policy. *Curriculum Journal*, *31*(3), 357-378.

Gauld, C., & Hukins, A. (1980). Scientific attitudes: a review. *Studies in Science Education*, *7*, 129-161.

Gericke, N., Högström, P., & Wallin, J. (2023). A systematic review of research on laboratory work in secondary school. *Studies in Science Education*, *59*(2), 245-285.

Holman, J. (2017). *Good Practical Science*. The Gatsby Charitable Foundation.

McAlinden, M., & Noyes, A. (2019). Assessing mathematics within advanced school science qualifications. *Assessment in Education: Principles, Policy & Practice*, *26*(3), 340-355.

Millar, R., & Abrahams, I. (2009). Practical work: making it more effective. *School Science Review*, *91*(334), 59-64.

Moore, A. M., Fairhurst, P., Bennett, J. M., Harrison, C., Correia, C. F., & Durk, J. (2023). Assessment and practical science: identifying generalizable characteristics of written assessments that reward and incentivise effective practices in practical science lessons. *International Journal of Science Education*. https://doi.org/10.1080/09500693.2023.2253366

Moore, A. M., Fairhurst, P., Correia, C. F., Harrison, C., & Bennett, J. M. (2020). Science practical work in a COVID-19 world: are teacher demonstrations, videos and textbooks effective replacements for hands-on practical activities? *School Science Review*, *102*(378), 7-12.

Wellcome Trust. (2020). *Young people's views on science education: Science Education Tracker 2019, Wave 2*. Wellcome Trust.

Prior Knowledge in Physics: The Distribution Within Different Subgroups of Students Before Attending Physics Minors

Kevin Schmitt, Verena Spatz

Technical University of Darmstadt, Germany

Against the background of continuously high dropout rates, research on the prior knowledge of students follows a long tradition, especially in mathematics and STEM. So far the focus has mostly been on major students. Therefore, we aim at examining the prior knowledge in physics of minor students in selected fundamental content areas (mechanics, electricity, optics). In a preceded pilot study, a test instrument, scaled according to the Rasch model, was developed. With the main study, put forward in this contribution, we now turn to describing the prior knowledge in physics of students in five different non-physics STEM degrees (biology, civil- and environmental engineering, chemistry, mechanical- and electrical-engineering,).

For this purpose, we have investigated person abilities of the students based on the Item-Response-theory (IRT) and examined the results for significance and calculated descriptive statistics. Our results indicate that the subgroups differ depending on factors such as the participated lecture, high school semester or gender.

We envision that these results will build a basis for more detailed analyses in the future in order to adjust each physics lecture addressee-oriented, according to minor students' particular weaknesses and strengths.

Keywords: prior knowledge, physics minors, rasch-analysis.

Introduction

For a long time, STEM faculties have suffered from low student retention in Germany (Heublein, Hutzsch, & Schmelzer, 2022). Therefore many support programmes were introduced, especially for university freshmen (Bausch et al., 2014; Schmitt & Spatz, 2020). It could already be shown that pre-courses in mathematics can have positive effects for non-major mathematics students on the study success, measured by exam results (Büchele, 2018). However, dropout rates are continuously high and support measures more important than ever, not only in mathematics but also other subjects, with the heterogeneity, measured on different factors, in the composition of the student body increasing (Buschhüter, Spoden, & Borowski, 2017). In order to offer promising support programmes, these measures must cover up the prior knowledge that students bring to the

lecture to identify and close existing knowledge gaps. Therefore, many research projects on the prior knowledge of students were carried out, mainly focusing on the influence of prior knowledge in mathematics on study success (Müller, 2018). Lately, some studies looked at the influence of prior knowledge in physics and showed that it is also a valid predictor for the study success of physics majors (Binder et al., 2019). The focus of this contribution, the group of physics minor students, also got attention in prior knowledge research in 2020 (Sternal & Walliser, 2020). Namely, a study about the prior knowledge of first-year STEM discipline students which was conducted in the time span of 2014 to 2019. The study shows that the overall prior knowledge stays at an almost constant level over the measured years with slightly shifted content-based strengths and weaknesses and a vast heterogeneity between the students. Therefore, Sternal and Walliser used a prior knowledge test which includes eight items covering thermodynamics, mechanics and electricity. For our study, these results don't serve the needs completely. Since physics minors are considered to be so-called "gatekeeper courses" (Sorge, Petersen, & Neumann, 2016), we try to get a deeper insight of the state of the physics prior knowledge of these students. A research gap still remains in case of an detailed overview of the quality of prior knowledge in different content areas which would be comparable to studies from Binder et al. (2021). The heterogeneity in prior knowledge of students partly results of physics being non-obligatory in obtaining university entry qualifications. Therefore, we try to focus on collecting information about which contents are lacking most in prior knowledge to address them subsequently to this study. In addition, the information can help lecturers to identify the needs of students and adjust learning activities based on the prior knowledge levels. Consequently, we designed a prior knowledge test for physics minor students which covers up a variety of physical contents comparable to the named studies.

Research Goal

Overall, our project was oriented on three coherent steps to approach the overarching research goal. The first step was the development of a prior knowledge test for physics minor students which only contains physics contents and different knowledge types, identified by Hailikari et al. (2007). The second step was to conduct the main study, including the usage of the developed instrument for all physics minor students at the Technical University of Darmstadt. The last step, which covers up the contents of this paper, is to examine and statistically analyse the prior knowledge in physics of university freshmen in non-physics STEM degrees. To this end, we also investigate differences between physics minors in various subgroups (biology, civil- and environmental engineering, chemistry, mechanical- and electrical-engineering students).

Method

Test instrument

Based on an existing instrument of Binder et al. (2019), we have developed a test instrument and scaled it according to the Rasch model in a preceded pilot study (Schmitt & Spatz, 2022). The pilot study is composed of two different measurement times. Therefore, we could identify good and bad fitting items and adjust them according to the needs of the subgroup after the first pilot study. In the second pilot study we verified the changes and that the test instrument can be used according to the quality criteria based on Item-Response-Theory.

The test splits into three parts, each covering one domain of knowledge (*knowledge of facts, knowledge of meaning, application of knowledge*) according to Hailikari's model (Hailikari, 2009). In the original knowledge model and the preceded studies of Binder et al. (2019), another knowledge type and therefor test section was used: *integration of knowledge*. Since it was shown that the test section had no incremental validity in predicting study success and the lack of practical implementation possibilities in our study frame, we excluded this type from our research. Each test part contains items out of the content areas of mechanics, electricity and optics; fundamental content areas which have revealed relevant in all assessed physics lectures for minor students. The item types differed for the three test sections and were chosen to fit the needs of the measured knowledge types at best. As a result the first part, *knowledge of facts* consists of multiple-choice questions with three wrong answers and one right answer. The second part *knowledge of meaning* is conducted by items which require free-text answers with a length of up to two sentences. In the third section *application of knowledge* is a drop-down format which asks students to choose the correct solution approach to calculation tasks with physical contexts.

The evaluation of the *knowledge of facts* and *application of knowledge* sections was conducted dichotomously ("right": 1 or "wrong": 0). In the *knowledge of meaning* sections, the answers allowed a more complex evaluation. Therefore, we created a rubric to measure the quality of the answers and control interrater reliability. Each question was rated by the core aspects of the right answer, e.g. for the item *"describe newton's first law"* we defined the core aspects to "*changing of the state of motion"* and *"external force".* Based on the quality and the description of the core aspect in the students answer, each aspect was rated by *0 – not contained, 0.5 – contained but wrong or incomplete* or *1 – contained correctly*. The sum of the rated aspects formed the overall item result.

Data acquisition

To gather our data this instrument has been used for a digital, online assessment as part of students' homework assignment in five physics lectures for minor

students (”physics for biology”: students’ first term of study, physics for electrical-engineering”: students’ first term of study, “physics for civil- and environmental engineering”: students’ second term of study, ”physics for chemistry”: students’ third term of study, ”physics for mechanical-engineering”: students’ third term of study) during two weeks at the beginning of the winter semester 2022/23 and summer semester 2023. The overall test duration was 45-60 minutes, including 43 questions distributed over the three test parts.

Overall, the sample consists of N = 702 data sets, of which at least N = 470 (335 male; 135 female) could be analysed. The other data sets were interrupted midway through and thus highly incomplete with many missing values. The sample sizes of the subgroups are distributed as follows: biology $N_{Bio} = 45$, civil- and environmental engineering $N_{C./E.} = 21$, chemistry $N_{Chem} = 23$, electrical-engineering $N_{El.Eng.} = 218$ and mechanical-engineering $N_{Mech.Eng.} = 163$.

Data analysis

The analysis has been carried out for each test part separately and as an overall result. We first used Rasch-analysis based on Item-Response-theory with the software “R” to calculate item and person parameters (item difficulties and person abilities). The parameters are measured as latent variables based on the responses of the students. We calculated person reliability as an IRT equivalent to the traditional test reliability such as Cronbach’s Alpha for each test section individually, showing good to very good internal consistency from *application of knowledge* (.71) and *knowledge of facts* (.78) to *knowledge of meaning* (.81). In addition the overall test shows very good reliability with a value of .89.

The rating concept of the second test section was analysed for interrater agreement by calculating Cohen’s Kappa for different pairs of raters (Cantor, 1996). The overall kappa value, calculated by the mean of the single kappa values of the items, is .86.

After Rasch-analysis, we analyse the person abilities of the different subgroups with traditional statistical methods, such as analysis of variances (ANOVA) including effect sizes conducted by partial η^2 (Okada, 2013) and post-hoc tests (Tukey HSD). As independent variable we used the person abilities and as dependent variables (with the corresponding values or clusters): the lecture (*Biology; C./E.-Engineering; Mechanical Engineering; Chemistry; Electrical Engineering)*, high school grade (*0.5 grade steps beginning at 4 and ending at 1)*, gender (*male; female; other – which isn’t represented)*, semester of matriculation (*less than 3; higher than 3 and less than 6; higher than 6 and less than 9)*. Different calculations of assumed ANOVA models, including some with interactions between independent variables, could show, that there are no significances in interactions and the best fitting model is the one excluding these.

Results

The results will be presented separately for each test section, differed by the participating lectures. Figure 1 shows the prior knowledge calculated as mean person ability where lower values indicate lower levels of prior knowledge and vice versa. A person ability of 0 can be interpreted as a person's estimate to answer a question correctly at 50%. For each lecture the mean values are shown in every test section.

In addition, the results of the ANOVAs for each test section are shown in table 1. We limit the report of the post-hoc test results to significant ones. All the results will be summed up since the ratio of the person abilities of the different lectures does not change across the test sections.

Figure 1. Grouped bar-charts for mean person abilities differed by the lecture addressed major for every knowledge type. *KoF:* knowledge of facts, *KoM:* knowledge of meaning and *AoK:* application of knowledge.

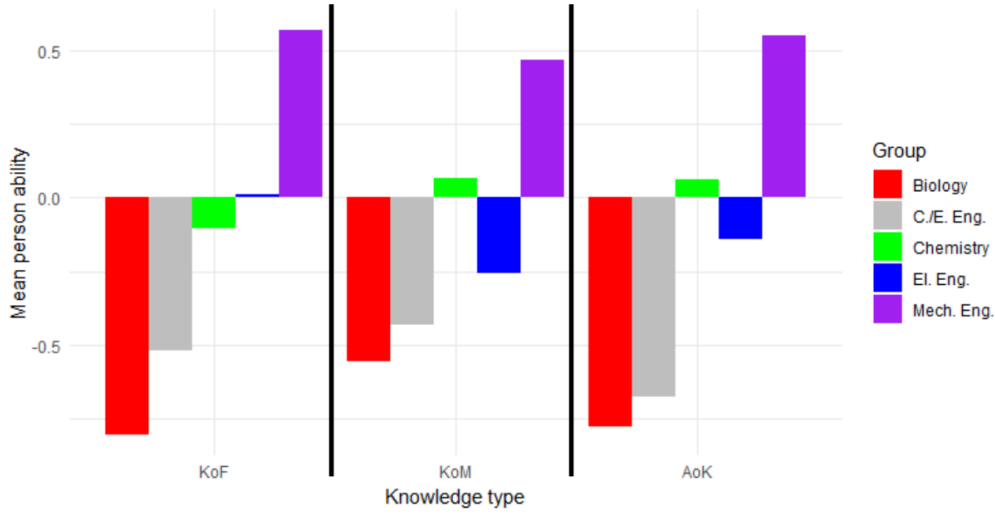


Table 1. F-values of ANOVA's and partial η^2 values as effect sizes for the different test sections (*$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$).

		Knowledge of facts		Knowledge of meaning		Application of Knowledge	
	df	**F-Value**	**Partial η^2**	**F-Value**	**Partial η^2**	**F-Value**	**Partial η^2**
Lecture	4	15.25***	**0.12**	14.37***	**0.11**	12.08***	**0.10**
High school grade	5	6.15***	**0.07**	8.62***	**0.09**	6.15***	**0.07**
Gender	1	22.43***	**0.05**	7.08*	**0.02**	23.05***	**0.05**
University semester	2	1.04	**0.005**	0.41	**0.002**	2.20	**0.01**
Residuals	445						

The bar chart in figure 1 indicates how the students of the different lectures performed in the tests compared to each other. Participants of the lectures "physics for biology" and "physics for civil- and environmental engineering" appear to have the least prior knowledge compared to the other groups (knowledge of facts: $M_{Bio} = -0.80$; SD =1.09, $M_{UBI} = -0.52$; SD = 0.98, knowledge of meaning: $M_{Bio} = 0.55$; SD = 0.97; $M_{UBI} = -0.43$, application of knowledge: $M_{Bio} = -0.78$; SD = 1.19, $M_{UBI} = -0.68$; SD = 1.31). If we take a closer look at the person abilities of these groups, it appears that the knowledge of meaning test section shows the best results compared to the other knowledge types. The highest performing subgroup of participants was the "physics for mechanical engineering" lecture (knowledge of facts: $M_{M.E.} = 0.57$; SD = 1.19, knowledge of meaning: $M_{M.E.} = 0.47$; SD = 1.02, application of knowledge: $M_{M.E.} = 0.55$; SD = 1.42). With a slightly lower mean value of person ability in the section knowledge of meaning compared to the other two knowledge types. Participants of the "physics for chemistry" as well as "physics for electrical engineering" lectures showed average mean person abilities at around 0 (knowledge of facts: $M_{Ch} = -0.10$; SD = 1.02, $M_{E.E.} = 0.01$; SD = 1.35, knowledge of meaning: $M_{Ch} = 0.07$; SD = 1.14, $M_{E.E.} = -0.26$; SD = 1.26, application of knowledge: $M_{Ch} = 0.06$; SD = 1.28, $M_{E.E.} = -0.14$; SD = 1.46). Whereas chemistry participants tend to have no ability differences between the knowledge types, electrical engineering participants perform worse in knowledge of meaning, especially compared to knowledge of facts.

For further analysis we will now report the results of the ANOVA shown in table 1 and post-hoc tests. As already discussed for the descriptive statistics of mean person abilities, the participated lecture shows a significant effect on the person ability for every knowledge type (knowledge of facts: F(4, 445) = 15.25***, knowledge of meaning: F(4, 445) = 14.37***, application of knowledge: F(4, 445) = 12.08***) with medium effect sizes (knowledge of facts: $\eta^2 = 0.12$,

knowledge of meaning: $\eta^2 = 0.11$, application of knowledge: $\eta^2 = 0.10$). Post-hoc comparisons could additionally show that the main differences are located between mechanical engineering students and all other subgroups, except for chemistry in the sections knowledge of meaning and application of knowledge. Additionally, there are significant differences between the groups of chemistry and electrical engineering in the knowledge of facts section and the group of biology students.

The variable *high school grade* showed comparable results in ANOVA: there are highly significant differences for every knowledge type (knowledge of facts: $F(5, 445) = 6.15^{***}$, knowledge of meaning: $F(5, 445) = 8.62^{***}$, application of knowledge: $F(5, 445) = 6.15^{***}$) also having medium effect sizes on the lower end (knowledge of facts: $\eta^2 = 0.07$, knowledge of meaning: $\eta^2 = 0.09$, application of knowledge: $\eta^2 = 0.07$). A closer look into post-hoc test results shows that the differences appear mostly between the group of the grade range *1-1.5 (best grade range)* in comparison to all other grade ranges. Between the other grade ranges, there appeared no significant differences in test results.

A comparison of the *gender* variable showed significant differences as well, especially for the knowledge of facts and application of knowledge sections (knowledge of facts: $F(1, 445) = 22.43^{***}$, knowledge of meaning: $F(1, 445) = 7.08^{*}$, application of knowledge: $F(1, 445) = 23.05^{***}$). However, the effect sizes were small for each test section for this variable (knowledge of facts: $\eta^2 = 0.05$, knowledge of meaning: $\eta^2 = 0.02$, application of knowledge: $\eta^2 = 0.05$). The results of the post-hoc tests indicate that male students had significant better test results than female students.

In the ANOVA, as well as the post-hoc test, the variable *high school semester* showed no significances at all which is an interesting finding and will be discussed in the upcoming section.

Discussion

The overall test results show significant differences between physics minors in various subgroups (biology, chemistry, mechanical- and electrical-engineering students). Thus, these significances are not surprising, as students' motives of choice for their degree might be closer related to physics for students of mechanical- engineering than for students of biology. Out of this perspective it seems interesting that students of the other engineering faculty perform much worse than the other subgroups. Reasons could be that the sample size of the group does not represent an adequate group of the students and the curriculum is less mathematics and science focussed than the mechanical engineering one. These results lead to the conclusion that the different subgroups need differently designed support measures to adapt the prior knowledge in physics of the students. In addition to the differences between the lectures, by analysing the

distribution functions we could also show that the mean person abilities of the individual subgroups of students in each lecture vary a lot as well. This makes support measures even more important since we aim to give students the chance to start with equal entry requirements when it comes to prior knowledge in physics lectures.

Out of this perspective to former research projects, the results for the variable *high school grade* replicate that there is an influence of the high school grade and the performance at university. However, it appears to differ in the way that the high school grade in the majority of studies tends to be the single best prediction factor for study success (Fleischer, Averbeck, Sumfleth, Leutner, & Brand, 2016; Sorge et al., 2016; Trapmann, Hell, Weigand, & Schuler, 2007). In fact, our study cannot directly be compared to the named studies but could lead to further examinations of the relationship between the success in the final exam of physics minor lectures and the prior knowledge assessed by our test instrument.

That the semester of matriculation showed no significant differences in the ANOVA when comparing all subgroups lead to additional analysis. Therefore, we conducted a further analysis of the individual subgroups of students in the different lectures with a view to the high school semester which also showed no significant differences in person abilities for this variable. A possible explanation for this result could be that there are cancelling effects of students participating in the physics minors in higher semesters because they have difficulties in their studies. Another explanation would be that already attended non-physics lectures in the curriculum don't contribute to physics prior knowledge at all, even in engineering courses.

As a comparison to other studies, we could show that there is a difference in the performance of male and female students in our sample group. This could also be shown by Whitcomb & Singh (2020) where they found a small effect in gender difference for physics major students.

Limitations and Outlook

We acknowledge that overall - for example - the different sample sizes in the subgroups currently limit the informative value of the findings. This accounts especially for the participants in the lectures "physics for chemistry" and "physics for civil- and environmental engineering" with a sample size of only $N_{Chem.} = 24$ and $N_{C./E.} = 21$. Another limitation of our study is related to the independent variable of the *semester of matriculation*. Since the subgroups of students already differ naturally by the implementation of the lecture in the corresponding major curriculum, this factor may be influenced by the factor *lecture* as well, even if an ANOVA including the interrelation between these two variables showed no significance.

As an outlook to upcoming analyses, we will examine the results in specific content areas such as mechanics, electricity and optics for the named dependent variables. This could contribute to a more detailed and faceted picture of the prior knowledge in physics of university freshmen in non-physics STEM degrees. We hope that such insights will build a basis to adjust each physics lecture for minor students independently based on their particular weaknesses and strengths in order to make them more addressee-oriented. As a direct reaction to the shown results, we started a support programme for physics minor students, which consists of an online course. This is offered in an online self-assessment format that can be completed before the lecture starts and covers all topics in school physics and some advanced topics.

References

Bausch, I., Biehler, R., Bruder, R., Fischer, P. R., Hochmuth, R., Koepf, W., . . . Wassong, T. (2014). *Mathematische Vor- und Brückenkurse*. Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-03065-0

Binder, T., Sandmann, A., Sures, B., Friege, G., Theyssen, H., & Schmiemann, P. (2019). Assessing prior knowledge types as predictors of academic achievement in the introductory phase of biology and physics study programmes using logistic regression. *International Journal of STEM Education*, *6*(1). https://doi.org/10.1186/s40594-019-0189-9

Binder, T., Waldeyer, J., & Schmiemann, P. (2021). Studienerfolg von Fachstudierenden im Anfangsstudium der Biologie. *Zeitschrift Für Didaktik Der Naturwissenschaften.* Advance online publication. https://doi.org/10.1007/s40573-021-00123-4

Büchele, S. (2018). *Bridging the gap: How effective are remedial math courses in Germany?* (MAGKS Joint Discussion Paper Series in Economics No. 25-2018). Retrieved from Marburg: Philipps-University Marburg, School of Business and Economics website: https://www.econstor.eu/handle/10419/200681

Buschhüter, D., Spoden, C., & Borowski, A. (2017). Physics knowledge of first semester physics students in Germany: a comparison of 1978 and 2013 cohorts. *International Journal of Science Education*, *39*(9), 1109–1132. https://doi.org/10.1080/09500693.2017.1318457

Cantor, A. B. (1996). Sample-size calculations for Cohen's kappa. *Psychological Methods*, *1*(2), 150–153. https://doi.org/10.1037/1082-989X.1.2.150

Fleischer, J., Averbeck, D., Sumfleth, E., Leutner, D., & Brand, M. (2016). Entwicklung und Vorhersage von Studienzufriedenheit in MINT-Fächern. In C. Maurer (Chair), *Jahrestagung.* Symposium conducted at the meeting of Gesellschaft für Didaktik der Chemie und Physik (GDCP), Zürich.

Hailikari, T. (2009). *Assessing University Students' Prior Knowledge.: Implications for Theory and Practice.* (Dissertation). University of Helsinki, Helsinki.

Hailikari, T., Nevgi, A., & Lindblom-Ylänne, S. (2007). Exploring Alternative Ways Of Assessing Prior Knowledge, Its Components And Their Relation To Student Achievement: A Mathematics Based Case Study. *Studies in Educational Evaluation*,

33(3-4), 320–337. https://doi.org/10.1016/j.stueduc.2007.07.007

Heublein, U., Hutzsch, C., & Schmelzer, R. (2022). *Die Entwicklung der Studienabbruchquoten in Deutschland*. Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW). https://doi.org/10.34878/2022.05.DZHW_BRIEF

Müller, J. (2018). *Studienerfolg im Fach Physik: Zusammenhang zwischen Modellierungskompetenz und Studienerfolg* (Dissertation). Universität Duisburg-Essen.

Okada, K. (2013). Is Omega Squared Less Biased? A Comparison of Three Major Effect Size Indices in One-Way Anova. *Behaviormetrika*, *40*(2), 129–147. https://doi.org/10.2333/bhmk.40.129

Schmitt, K., & Spatz, V. (2020). *Digital Support for physics service lectures*. Zenodo. https://doi.org/10.5281/ZENODO.3945813

Schmitt, K., & Spatz, V. (2022). Physikalisches Vorwissen in Physik-Nebenfachveranstaltungen. In H. Grötzebauch (Ed.), *PhyDid B, Didaktik der Physik Beiträge zur virtuelle DPG-Frühjahrstagung 2022*.

Sorge, S., Petersen, S., & Neumann, K. (2016). Die Bedeutung der Studierfähigkeit für den Studienerfolg im 1. Semester in Physik. *Zeitschrift Für Didaktik Der Naturwissenschaften*, *22*(1), 165–180. https://doi.org/10.1007/s40573-016-0048-x

Sternal, O., & Walliser, N.-O. (2020). Physik-Vorkenntnisse von Studienanfängerinnen und –anfängern in MINT-Fächern. *Die Hochschullehre*. (6), 103–118.

Trapmann, S., Hell, B., Weigand, S., & Schuler, H. (2007). Die Validität von Schulnoten zur Vorhersage des Studienerfolgs - eine Metaanalyse. *Zeitschrift Für Pädagogische Psychologie*, *21*(1), 11–27. https://doi.org/10.1024/1010-0652.21.1.11

Whitcomb, K. M., & Singh, C. (2020). For physics majors, gender differences in introductory physics do not inform future physics performance. *European Journal of Physics*, *41*(6), 65701. https://doi.org/10.1088/1361-6404/ab9f1d



Strand 17

Science Education in Primary and Pre-primary Learning Contexts

Bodil Sundberg & Christina Siry

Strand Chairs/Section Editors

Foreword

We are pleased to present Chapter 17 of the ESERA 2023 Conference Proceedings, titled Science Education in Primary and Pre-primary Learning Contexts. This chapter features six proceedings that focus on the foundational stages of science education, where young learners first encounter the wonders of the natural world and begin to develop their scientific understanding.

Early science education plays a crucial role in shaping students' attitudes towards science and laying the groundwork for future learning. The studies compiled in this chapter explore various aspects of science teaching and learning within primary and pre-primary contexts, providing valuable insights into how young children engage with scientific concepts and develop inquiry skills.

The research presented here examines a range of topics, including effective instructional strategies for young learners, the role of play and exploration in science learning, and the integration of science with other areas of the curriculum. Additionally, the proceedings highlight the importance of fostering a positive and supportive learning environment that encourages curiosity, creativity, and critical thinking.

Understanding how children learn science at these early stages is essential for creating engaging and developmentally appropriate educational experiences. The findings shared in this chapter contribute to our knowledge of how best to support young learners in building a solid foundation in science, addressing both the challenges and opportunities inherent in early science education.

We extend our gratitude to the authors and reviewers who have contributed to this chapter. Their work underscores the significance of high-quality science education from the earliest years and offers practical insights that can inform and enhance teaching practices in primary and pre-primary settings.

As you explore the studies presented in this chapter, we hope you gain a deeper appreciation for the complexities of science education in early learning contexts and find inspiration for fostering a lifelong love of science in young learners.

Are Games Effective to Learn and Think about the Nature of Science in Primary Education?

Maria-Antonia Manassero-Mas, Ángel Vázquez-Alonso

University of the Balearic Islands, Spain

The literature acknowledges that the nature of science (NoS) is difficult to learn due to its metacognitive character, the need to master thinking skills, poor teacher training, and lack of resources, which are especially acute in primary education. This paper attempts to overcome these difficulties, presenting four epistemic games as educational resources that gamify scientific practices to teach NoS issues in primary classrooms. The games, their didactic materials, the explicit and reflective pedagogic procedures, and the scientific thinking skills are described. An attitudinal test on the image of science and technology (S&T) assesses the effectiveness of games through a quasi-experimental pre-post-test research design. The results show that students highly agree with the importance of S&T and its capacity to provide better opportunities, and a majority disagrees with trusting scientists. The games improve the overall image of S&T, yet the changes are significant in some specific aspects, such as solving nearly all problems and curing diseases. Gender analysis finds that girls tend to increase their agreement with the traits of beneficial and harmful effects, trust in scientists, and better future opportunities, whereas boys increase their mean scores on solving nearly all problems and curing diseases. Finally, the achievements of games to teach NoS and develop thinking, the consequences for learning, the limitations, and the flexibility and sustainability as prospective features of games are discussed.

Keywords: Nature of science, serious learning games, science-related attitudes

Introduction

This study focuses on the nature of science (NoS), a set of complex, multifaceted, interdisciplinary, evolutionary, and changing pieces of knowledge, procedures, and values that scientists use in scientific practice to validate scientific knowledge and intervene in society (Erduran & Dagher, 2014; Manassero-Mas & Vázquez-Alonso, 2019). Properly speaking, NoS should be called the nature of scientific knowledge because this label better reveals its meta-knowledge content or the knowledge about knowledge (what do we know?) and about knowing (how do we know what we know?). Furthermore, it justifies the convergence of NoS with the epistemic knowledge used in the educational literature (Greene et al., 2016).

The metacognitive complexity of NoS determines its difficulties for teaching and learning. The literature focuses on the high cognitive demand, teachers' insufficient understanding, and the scarcity of teaching resources. All three factors are more severe in the Spanish context due to the inflation of educational laws over the years, which decreed school curricula that ignored explicit teaching and teacher training about NoS. Consequently, textbooks and teaching were deprived of NoS as a valuable science education innovation. Furthermore, Spanish primary teachers' training in general science, and particularly in NoS, is traditionally poor, which adds to the fact that most Spanish primary schools have recently given up their laboratories, thus losing a pivotal resource to link NoS and inquiry learning (Vázquez-Alonso & Manassero-Mas, 2017; García-Carmona, 2022). Thus, the main aim of this study is to cope with this unfavorable context to test the opportunities for effective NoS teaching and learning.

This study advocates for serious games, which have been previously explored and selected, as the primary pedagogical tool to overcome these difficulties because serious games may attain difficult learning objectives through their enjoyable, motivating, and affordable ability to engage students in learning during analogical scientific practices (Vázquez-Alonso & Manassero-Mas, 2016). Epistemic games are serious games focused on solving realistic, authentic, open-ended, and multidisciplinary problems within the epistemic framework of a particular community of practice. Games that simulate problem-solving, as experienced in scientific practice, can raise the NoS issues (Oubahssi et al., 2020). Furthermore, several studies suggest that serious learning-oriented games are more effective than conventional instruction in producing behavioral, cognitive, motivational, affective, and sociocultural changes (Clark et al., 2016; Plass et al., 2019).

The pedagogical approach to applying games to teach NoS in primary education draws on the factors impacting learning from the visible learning theory (Hattie, 2018). In short, the influential factors (Cohen's $d > .60$) involve general learning strategies (effort—to win games—, deliberate practice—continuous practical activity—, and summarization—checking the game evolution); metacognitive and self-regulated learning strategies (organization, evaluation, reflection, help-seeking and transfer); feedback strategies (response to actions, classroom discussion, feedback); goal-oriented learning strategies (planning and prediction, appropriately challenging goals, cognitive task analysis, discriminating goals from non-goals); instructional strategies (problem-solving, reciprocal peer teaching, jigsaw method); personal abilities (deep motivation, self-efficacy, and field independence); and curricular strategies (creativity and conceptual change).

Finally, gender is an extremely influential variable in cognitive, procedural, and attitudinal aspects related to science education. In developed countries, women tend to obtain lower knowledge scores than men, suffer discrimination

in laboratories and science classes, and show general attitudes of lower interest toward science topics, which produces an under-representation of women in many studies in the STEM field, especially in technologies and engineering, recognized as the gender gap (Aguillon et al., 2020; Quinn et al., 2020). Therefore, equity aimed at eradicating the male stereotype of science and promoting the active inclusion of women is a universal goal at all levels of science education and is also assumed herein as a core goal of games (Cheryan et al., 2017).

This paper presents four epistemic games on NoS topics within the pedagogical framework of explicit and reflective teaching of NoS, which help students learn NoS effectively. The games develop and practice scientific thinking skills and provide professional development to teachers through planning their NoS lessons with games in primary education. However, primary education is the least frequently used educational research scenario for NoS teaching and learning due to students' low cognitive developmental level and the primary teachers' scarce training in NoS (Cofré et al., 2019). Thus, this study also aims to fill in this relative gap of NoS research in primary education.

This paper aims to show the four epistemic games and test them as effective tools for scientific thinking-based NoS learning by improving students' S&T image and contributing to gender equity. The specific research questions are: What is the primary students' S&T image? Are games effective in improving S&T's image? Which specific S&T image traits are improved by games? How do games impact boys and girls?

Methods and Materials

The research methodology used a quasi-experimental pre-post-test design about teachers' application of games to students within the classrooms. The test on the image of S&T was applied before (pre-test) and after (post-test) the classroom game experience. The post-test took place at the end of the school year, several weeks after the end of the last game application.

Materials

As didactic resources, the four serious games satisfy the basic criteria for effective NoS teaching: explicitly presenting the NoS topics and stimulating thoughtful reflections among peers about epistemic (created ideas, hypotheses, predictions, argumentation, etc.) and social (cooperation, competition, discussion, attitudes) issues. Table 1 presents each game synthetically.

The games include the resources needed to play each game (pieces, cubes, decks of cards, boxes), the didactic documents for guiding teaching and learning. The didactic documents comprise the students' organizing guides, a short reading on a case drawn from the history of science, the teaching-learning sequence (lesson

plan), and a video in which a cutting-edge female scientist explains her area of scientific practice (Table 1).

Table 1. Summary of the materials and nature of science ideas involved in each game.

Game / Challenge	**Resources**			Nature of Science Topics and Skills	
	Materials	Readings	Didactics	Specifics	Transversals
Tangram or puzzle					
Compose a square with 4 pieces; then, do the same with 5 pieces	5 tangram pieces	Where do living beings come from?	Students' organizers TLS*	Tentativeness (provisional and fallible knowledge) Change and paradigm shift	Observing Data and evidence coordination through argumentation Creating hypothesis Reasoning Arguing to convince others Cooperating to deal with ideas and concepts Sharing contents and arguments with others Practicing / participating in discussions Scientific attitudes (curiosity, openness, skepticism...)
Cubes or dices					
What is on the hidden face of the cube?	Card Cubes (or Paper Cubes)	A New Planet	Students' organizers TLS*	Identifying patterns Arguing with evidence Elaborating explanations Making predictions	
Cards					
Which law are the cards following?	Decks of cards	A law for navigation	Students' organizers TLS*	Identifying complex patterns Arguing with evidence Elaborating explanations Making predictions	
Mystery or black boxes					
What is inside the box?	Closed box full of various objects	Classification of living beings Opening the Mystery Box of Matter	Students' organizers TLS*	Arguing with evidence Elaborating explanations Making predictions Practicing inquiry skills (measures, classification, design, checks…)	

*TLS: Teaching-learning sequence of the game.

Research instrument

Before (pre-test) and after (post-test) playing the games in the classroom, the students answer a 7-item attitudinal test on the image of S&T, which is adapted for primary students from the ROSES questionnaire (Jidesjö et al., 2020). Each item displays a short, simple, and primarily positively-worded sentence (e.g., S&T are important for society). The students codify their responses on a 4-point Likert scale (1= disagree, 4 = agree), such that the higher the score, the better the S&T image.

Participants

The participants in this experience of scientific practice with games are 12 teachers from five different schools who taught the games to their sixth-grade primary students (165 students; 78 girls and 88 boys) within their natural class groups.

Procedures

The researchers to teachers presented and taught the didactic documents of games and then the teachers appropriate the materials to accordingly prepare, plan, teach, and direct the game activities for their students. Overall, these in-service teaching activities are expected to achieve professional development on NoS for participant primary teachers.

Teachers did not report significant incidents during the game activities, except that the students demanded playing more time because they enjoyed the games. Teachers spent about 6 hours on game lessons on average, and the black box game was only conducted in one school due to time constraints.

All games pose a challenge that must be met by small groups of students (escape games), which triggers cooperative group work in seeking the response to the games' challenges. These procedures make the game experience similar to a group problem-solving and cooperative project through scientific thinking skills, with which the participant teachers were familiar. Furthermore, these procedures align with the model of validation of scientific knowledge in scientific practice through the observation, argumentation, cooperation, and competition among peers (Capraro et al., 2013; Tena & Couso, 2023).

A golden rule was established to stimulate reflection: all students' claims during the game must be consistent with observations, evidence, or available knowledge, and this consistency was determined by cooperatively using scientific thinking skills (evidence-checking). This golden rule means that to be accepted, the students' proposals during the game activities (ideas, hypotheses, predictions, etc.) should always meet at least one of the following conditions: students should accompany the idea with a justification, argument, or reasoning,

or they should support the idea's consistency with the available evidence.

The teachers applied the assessment questionnaire to their students in digital format, following its protocols and rules to diagnose the long-term effects of games. Gender differences were assessed using the effect-size statistic (Cohen's d), which measures differences in standard deviation units and is more practical than the probability of statistical significance (p) because it reports the magnitude of the difference (small: $d < .20$).

Results

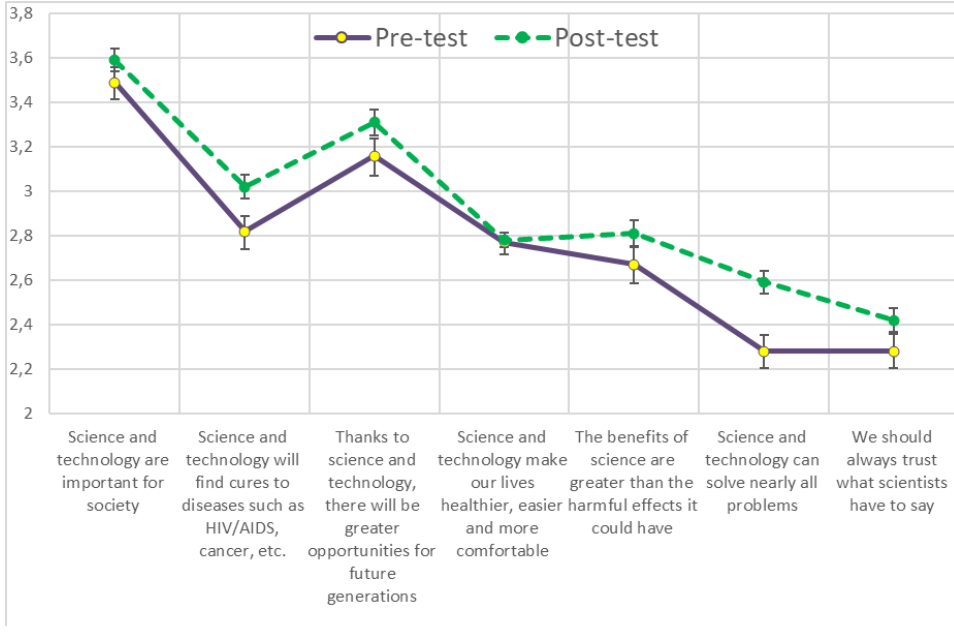
The pre-test results before playing the game (brown bold line of Figure 1) in the test on the S&T image show that the students significantly agreed (mean > 3) with two items: "S&T are important for society," and "provide greater opportunities for future generations." On the other hand, "trust in scientists" obtained the lowest mean agreement score.

The results after the students played the games in their classrooms (dotted green line of Figure 1) show that students significantly agreed (mean > 3) with the following three items: "S&T are important for society," "they will find cures to diseases," and "provide greater opportunities for future generations." Again, "trust in scientists" obtained the lowest mean agreement score.

Impact of games on S&T image

The effectivity of the game experience in changing students' attitudes toward S&T was empirically and quantitatively assessed by the pre-post differences in the seven items of the S&T image (Figure 1). The magnitude of the change potentially attributable to the games was assessed through the statistical significance and the effect size (Cohen's d) of the differences between the pre-test and post-test weighted mean scores (Figure 1). First, we highlight the main overall result conveyed by Figure 1: all post-test scores of the seven items about the image of S&T were higher than the mean pre-test scores, showing that games may improve the image of S&T. Second, the most significant improvement appears on the S&T item "can solve nearly all problems" ($p = .005$, $d = 0.34$). The S&T item "will cure diseases" ($p < .05$, $d = 0.22$) showed the second most significant improvement.

Figure 1. Pre-test and post-tests weighted mean scores in the items of the image of S&T and their standard error bars.



All in all, these results confirm the potential positive effect of games for improving the image of S&T, which includes all the basic aspects of NoS and some significant ones. Thus, the conclusion is that games may contribute to improving the students' NoS understanding.

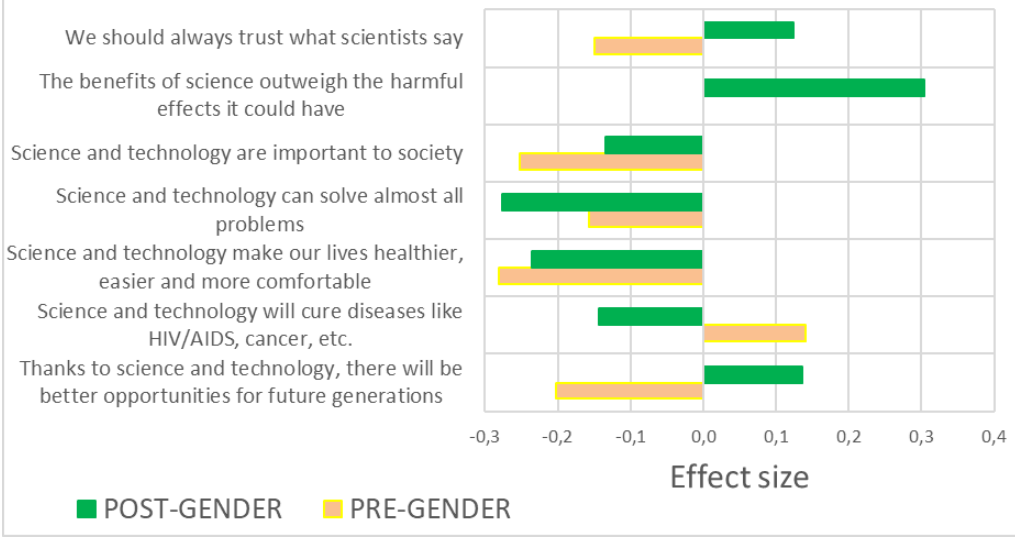
Girls-Boys attitudinal differences

The attitudinal differences between boys and girls were analyzed through several comparisons. First, gender differences were analyzed separately at the two moments considered in this study, namely, before and after playing the games (Figure 2). Second, the pre-post differences were separately analyzed, on the one hand, for the group of boys and, on the other hand, for the group of girls (Figure 3). To simplify the description, the effect size of the differences (Cohen's d) was used to quantify the amount of the differences; as a complement, when Cohen's $d > 0.20$, the differences were considered statistically significant ($p < .05$).

The overall attitudinal differences between girls and boys before the game experience (pre-test) showed that boys tended to obtain higher attitude scores than girls, as five items presented negative differences (boys' mean scores were higher than girls' mean scores, $d < 0$), whereas girls scored higher than boys ($d > 0$) only in one item ("S&T will cure diseases"). However, the differences were nonsignificant, as the absolute scores of the effect size of the differences for all items (without considering the positive or negative sign) were less than .30.

The pattern of gender differences after the game experience (post-test) continued to show small gender differences. Only one item ("the benefits of science outweigh the harmful effects") presented large differences, where it is noteworthy that girls scored significantly higher than boys (d > .30). Further, the overall comparison between girls and boys was quite balanced, as girls scored higher than boys on three items, whereas boys scored higher on four items. This ratio (three/four) slightly reverses the global pre-test pattern of the differences, which was favorable to boys (one/five).

Figure 2. Effect size of gender differences before (pre) and after (post) playing the games.



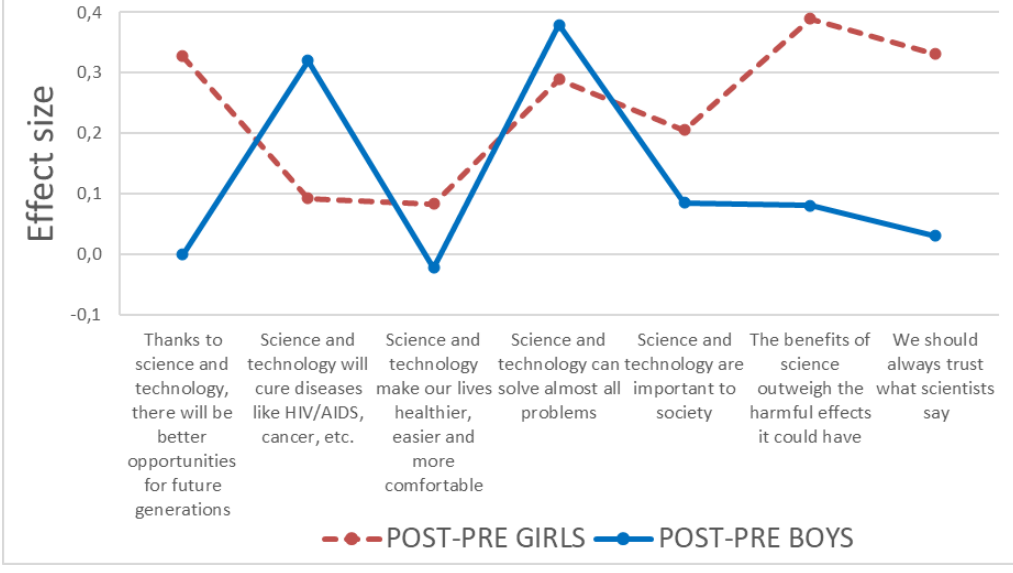
In sum, none of the attitudinal differences between girls and boys before and after playing the game were significant or relevant (d < .40). However, the overall qualitative and quantitative analysis indicated that girls improved their attitudes relatively more than boys, as shown by the increase in the number of items with differences in favor of the girls at post-test and the girls' higher final magnitudes of the differences (d) in several issues. This suggests that games may act by reversing the initial pattern of gender differences favorable to boys into a balanced profile between girls and boys, apparently benefiting girls over boys.

To confirm the former trend of attitudinal change, we computed the pre-post differences within each gender, analyzing them separately for the boys and the girls (Figure 3). This analysis evaluates the changes within each gender independently; that is, the magnitude of the impact of games for boys and for girls.

Girls improved their mean scores across all seven S&T items, which means that games may increase girls' agreement scores across all items. Moreover, girls displayed the most significant improvement ($d > .30$) on three items: "the benefits are greater than the harmful effects," "always trust what scientists say," and "thanks to S&T, there will be better opportunities for future generations." Boys' scores did not improve two of these items (opportunities and better life) and only improved their mean scores on five S&T items. Further, the two items where boys achieved the highest improvement ($d > .30$) were not the same items where the girls improved: "S&T can solve nearly all problems" and "S&T will cure diseases."

Overall, these results present a slightly different picture for girls and boys about the potential influence of games, as indicated by the former analysis. On the one hand, girls showed a generalized improvement across all S&T items, and no item decreased its scores, whereas boys' improvement was partial, as two items did not improve. The girls' quantitative improvement was greater than the boys', as the girls displayed larger effect sizes than the boys across five S&T items. On the other hand, the items showing the greatest improvement were different for girls (benefits, trust in scientists and opportunities) and boys (solving all problems and curing diseases). All in all, although the gender differences and the potential change due to the games were not quantitatively large, both qualities indicate a higher impact of games on girls than on boys, which is a noteworthy qualitative trend.

Figure 3. Effect size of the differences between pre-test and post-test mean scores of the seven items for the boys and the girls.



Negative pre-post attitudinal changes always represent some weaknesses of the educational process, as they pinpoint inefficiencies or undesired effects: a wake-up call to solve them through appropriate educational interventions, which should consider gender differences. On the one hand, the cross-analysis according to the sign of pre-post change for boys and girls (Figure 3) showed that five questions presented the same positive sign of attitudinal change for boys and girls. These items require educational care that considers the relative differences for boys and girls, namely "curing diseases" for girls, and "opportunities," "healthier life," "importance," "benefits," and "trust in scientists" for boys.

On the other hand, the issues conveyed by other items that show large changes between boys and girls do not just diagnose the potential weaknesses of science education. As they add an important gender difference, the appropriate educational care to address them should consider these differences in order to design inclusive lesson plans that are sensitive to the different signs of boys' and girls' attitudes.

Discussion and Conclusions

This study proposes an innovative way of teaching NoS in primary education through four epistemic serious games that provide a powerful, simple, and explicit analogy of scientific practice when students meet the game challenges cooperatively. The main contributions to the NoS field are using the games and their materials as a way to overcome the difficulties of learning NoS in primary education: lack of resources (the games develop new resources), scarce teacher training (the games train teachers to teach NoS), students' low cognitive development (modest NoS ideas and goals) and little research on the least frequent and most challenging scenario (primary education) for teaching NoS (Cofré et al., 2019; Tena & Couso, 2023).

The learning situations posed by the games to teach NoS apply an explicit and reflective methodology as consensually required by most NoS literature (Lederman, 2007). Overall, the games trigger high motivation to win the game and overcome the pedagogical barriers to build significant NoS learning naturally and playfully (Manassero-Mas & Vázquez-Alonso, 2020; Khishfe, 2020). The engaging activities of solving each game's challenge provide continuous opportunities for students to practice scientific thinking skills (creating hypotheses, arguing based on evidence, explaining and predicting observations, investigating, etc.), social skills (sharing, communicating, convincing, cooperating on projects, ideas, and concepts), values (curiosity, openness, skepticism, responsibility), and epistemic skills (evidence-reasoning coordination) (Erduran & Dagher, 2014).

The researchers helped the teachers to understand and appropriate the materials of the games to elaborate their own didactic transposition for their classrooms,

and subsequently to assess the games' NoS goals with a short attitudinal questionnaire on the students' image of S&T under a quasi-experimental design. The games adapted easily to students' cognitive development, teachers' training, and the varying needs of primary students due to their flexibility. Their simplicity ensures their sustainability to consolidate or extend the teaching of NoS topics within primary schools (Manassero-Mas & Vázquez-Alonso, 2020).

The quantitative and qualitative analyses display relevant findings to understand primary students' NoS learning. First, in reply to the first research question (S&T image), primary students mainly agree with some specific aspects, such as the general importance of S&T, and S&T's power to create better future opportunities, but they disagree about trusting scientists and S&T's power to solve problems. Second, the games are somewhat effective in improving students' NoS understanding because they improved their average scores in seven sentences of the questionnaire after playing the games, even though the quantitative differences are modest. The answer to the second research question (improvement of S&T's image) is deemed positive, although only a short time was devoted to playing games in the classroom: the games seem effective in fostering students' more favorable ideas about S&T because the students' final profile improved over the previous one. Third, the idea that S&T can solve nearly all problems showed the largest significant and relevant improvement, followed by trust in scientists and the perception of better opportunities, which also increased their final scores significantly. Thus, these ideas make up the answer to the third research question (traits improved by games).

The games also display some trends related to gender and the assessment items. The main finding shows the reduction of gender differences after playing the games (impact on boys and girls). The somewhat masculine pattern of gender differences before playing the games becomes a slightly feminine pattern at the end, when boys' and girls' attitudes are more similar than before playing the games. Overall, the girls' image of S&T improves much more than the boys' image, so the answer to the fourth research question identifies games as a friendly activity for girls, fostering their NoS understanding more than the boys'. This finding is worth highlighting, considering that many science education activities, especially laboratory practices, have been accused of being unfriendly to girls (Quinn et al., 2020).

The NoS traits identified as obtaining minority agreement scores indicate further opportunities to improve science education. For example, trust in scientists obtained overall low agreement, which is a wake-up call to elaborate this issue further within science education. This topic reflects the low status of NoS understanding, also showing its present paramount importance within the framework of the social controversies triggered by fake news, artificial intelligence, and the reliability of information sources. Furthermore, the traits in

which girls and boys obtain opposite scores (i.e., opportunities, advantages and disadvantages, trust in scientists) pose an additional controversy that complicates their improvement. The gender differences observed in this analysis require gender-inclusive educational pedagogies and adapted lesson plans so that both girls and boys can achieve the goal. In sum, the worsened or unchanged NoS traits suggest specific topics to improve primary science education, crucial to fostering NoS understanding and students' long-term future STEM participation (Aguillon et al., 2020; Quinn et al., 2020).

As mentioned above, the game experience involves many impact variables of visible learning (Hattie, 2018), but it also presents some limitations. First, the short journal space restricts the results presented herein, whose natural complement would be to extend the analysis to further items and documents of the experience. Likewise, the short time devoted to the games in the classrooms does not foster expectations of significant effects, although the findings are worthwhile contributions. Second, the NoS ideas presented herein are quite simple because greater epistemic depth would have been inappppropriate for primary students. Third, the games' analogical nature regarding scientific practice is compensated by providing students with readings about real examples from the history of science. Thus, this scaffolds and links each game to the curricular concepts and to thinking and learning about scientific practice (Fouad et al., 2015). Fourth, there was no control group because the researchers decided to include all the students in each school in the game experience. Fifth, although the sample of participants is quite diverse, its small size calls for new replications with larger samples and questionnaires to support the generalization of the results.

Finally, the game experience involves some noteworthy general implications for education. First, the games are cheap, flexible, and self-sustainable educational resources, which predict multiple didactic functionalities, such as their easy integration and adaption to school programs and grades, following the principles of open educational resources. Second, primary teachers receive easy and valuable in-service training to teach NoS and scientific thinking by studying the game materials, developing their specific lesson plans, and applying the game activities. Third, the teachers observed the students' impulsiveness to seek immediate rewards (win the game quickly) as a significant obstacle to learning NoS and scientific thinking effectively, as impulsiveness inhibits the necessary slow reflection. Overall, learning NoS and practicing scientific thinking provide primary students and teachers with some key skills to educate citizenship and character, which are acknowledged "vaccinations" against the current epidemic of information, misinformation, denialism, and pseudoscience (National Literacy Trust, 2018).

Acknowledgment

This research was funded by grant FCT-20-15728 of the Spanish Foundation for Science and Technology (FECYT).

References

Aguillon, S. M., Siegmund, G. F., Petipas, R. H., Drake, A. G., Cotner, S., & Ballen, C. J. (2020). Gender differences in student participation in an active-learning classroom. *CBE Life Sciences Education*, *19*(2), 12. https://doi.org/10.1187/cbe.19-03-0048

Capraro, R. M., Capraro, M. M., & Morgan, J. R. (2013). *STEM project-based learning*. Sense Publishers.

Cheryan, S., Ziegler, S.A., Montoya, A.K. & Jiang. L. (2017). Why are some STEM fields more gender-balanced than others? *Psychological Bulletin, 143*(1), 1-35. http://dx.doi.org/10.1037/bul0000052.

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research, 86*, 79–122. https://doi.org/10.3102/0034654315582

Cofré, H., Nuñez, P., Santibáñez, D., Pavez, J. M., Valencia, M., & Vergara, C. (2019). A critical review of students' and teachers' understandings of nature of science. *Science & Education, 28*, 205-248. https://doi.org/10.1080/09500693.2022.2152294

Erduran, S., & Dagher, Z. R. (Eds.) (2014). Reconceptualizing the nature of science for science education. *Scientific knowledge, practices and other family categories*. Springer. https://doi.org/10.1007/978-94-017-9057-4

Fouad, K. E., Masters, H., & Akerson, V. L. (2015). Using history of science to teach nature of science to elementary students. *Science & Education*, *24*, 1103–1140. https://doi.org/10.1007/S11191-015-9783-5

Hattie, J. (2018). *Visible learning, the sequel: A synthesis of over 2,100 meta-analyses relating to achievement*. Routledge. https://www.routledge.com/Visible-Learning-The-Sequel-A-Synthesis-of-Over-2100-Meta-Analyses-Relating/Hattie/p/book/9781032462035

Jidesjö, A., Oskarsson, M. & Westman, A-K. (2020). ROSES handbook: Introduction, guidelines and underlying ideas. *Utbildningsvetenskapliga studier 2020*:1. Mid Sweden University. http://www.miun.se/roses

García-Carmona, A. (2022). La comprensión de aspectos epistémicos de la naturaleza de la ciencia en el nuevo currículo de Educación Secundaria Obligatoria, tras la LOMLOE [Understanding epistemic aspects of the nature of science in Spain's new curriculum for compulsory-secondary education since the LOMLOE law]. *Revista Española de Pedagogía, 80* (283), 433-450. https://doi.org/10.22550/REP80-3-2022-01

Greene, J. A., Sandoval, W. A., & Bråten, I. (2016). *Handbook of epistemic cognition*. Routledge. https://doi.org/10.4324/9781315795225

Khishfe, R. (2020). Explicit Instruction and Student Learning of Argumentation and Nature of Science. *Journal of Science Teacher Education, 32*, 325-349. https://doi.org/10.1080/1046560X.2020.1822652

Lederman, N. G. (2007). Nature of science: Past, present, and future. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 831–879). Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203097267

Manassero-Mas, M.A., & Vázquez-Alonso, A. (2019). Conceptualization and taxonomy to structure knowledge about science. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias, 16*, 3104. https://doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2019.v16.i3.3104

Manassero-Mas, M. A., & Vázquez-Alonso, Á. (2020). Scientific thinking and critical thinking: Transversal competencies for learning. *Indagatio, 12*, 401-419. https://doi.org/10.34624%2Fid.v12i4.21808

National Literacy Trust (2018). *Fake news and critical literacy*. National Literacy Trust. literacytrust.org.uk/fake-news-resources.

Oubahssi, L., Piau-Toffolon, C., Loup, G., & Sanchez, É. (2020). From design to management of digital epistemic games. *International Journal of Serious Games*, *7*, 23–46. https://doi.org/10.17083/ijsg.v7i1.336

Plass, J. L., Mayer, R. E., & Homer, B. D. (2019). *Handbook of game-based learning*. MIT Press.

Quinn, K. N., Kelley, M. M., Mcgill, K. L., Smith, E. M., Whipps, Z., & Holmes, N. G. (2020). Group roles in unstructured labs show inequitable gender divide. *Physical Review Physics Education Research*, *16*. https://link.aps.org/doi/10.1103/PhysRevPhysEducRes.16.010129

Tena, E., & Couso, D. (2023). El diseño de preguntas investigables en el ciclo superior de primaria [The design of researchable questions in the primary upper cycle]. *Enseñanza de las Ciencias, 41*(1),101-123. https://doi.org/10.5565/rev/ensciencias.5573

Vázquez-Alonso, A., & Manassero-Mas, M.A. (2016). Games to teach the nature of scientific and technological knowledge. *Educar, 53*, 149-170. https://doi.org/10.5565/rev/educar.839

Vázquez-Alonso, Á., & Manassero-Mas, M.A. (2017). Science and technology contents in the new secondary education core curricula. *Revista de Currículum y Formación de Profesorado, 21*, 294–312. https://recyt.fecyt.es/index.php/profesorado/article/view/58064

Consideration of the Impact of a Professional Learning Program to Support the Teaching and Assessment of Practical Primary Science in England

Sarah Earle

Bath Spa University, UK

In England, primary science is often struggling for time in an increasingly crowded curriculum and schools are requesting stronger evidence of impact to justify the time and money that could be spent on any professional learning program. This means that programs like the one described in this paper need to be clearer than ever about their proposed impact on teacher practice. The Teacher Assessment in Primary Science (TAPS) Focused Assessment approach proposes that one focus is selected within the context of a whole inquiry, to support teaching and assessment. A three-day Focus4TAPS professional learning program has been developed to support teachers to implement this approach in primary science. The course has been delivered face-to-face and online. It has been the subject of a large-scale Education Endowment Foundation (EEF) randomized control trial in England. The EEF independent evaluation found a statistically significant impact on pupil learning. This paper will explore teacher survey data from the EEF evaluation, together with qualitative surveys completed by the course attendees during the program, to consider the process of implementation and described impact on practice. It is proposed that the TAPS Focused Assessment approach provides a manageable way for teaching and assessment of primary science inquiry, but that a sustained professional development program is necessary to support teacher understanding and formative use of assessment information gained from interaction with the children. This study provides insights that can inform the implementation of a similar approach in other contexts.

Keywords: Primary science, professional learning, teacher assessment

Introduction

The landscape of primary science in England is a mixed bag, with many organizations offering support, but schools struggling for curriculum time and primary science lacks status. The accountability structures of primary school league tables are based on results in English and mathematics, meaning that while on paper a 'core' subject, science is often a poor relative. With science organizations proposing that there should be two hours per week of science (ASE, 2020), the Wellcome Trust found that even schools who were keen enough

to reply to a survey, were only averaging 1.4 hours per week (Wellcome, 2017) and this issue has been exacerbated by the pandemic and a push for 'Covid catch-up'. Bianchi et al. (2021) described a range of issues in primary science practice, including lack of purpose in primary science activities arising from insecure teacher subject knowledge. A wide range of organizations continue to offer support and guidance for primary science in England (e.g. Primary Science Quality Mark, Ogden Trust partnerships, STEM Learning courses). One area with growing interest is for the development of 'primary science capital', supporting the children to see that science is 'for them', with teaching adapted, for example, to link to their interests (Archer et al., 2010; Nag Chowdhuri et al., 2021). However, with limited time, money and available supply cover, schools (and government) require strong evidence of impact to be convinced that professional learning programs for primary science are worth prioritizing.

The Focus for Teacher Assessment in Primary Science (Focus4TAPS) professional learning program developed out of the TAPS project, funded by the Primary Science Teaching Trust (PSTT) from 2013-22. The project took a Design-Based Research approach, developing theoretical and practical products through iterative cycles of collaboration with teachers and schools across the UK (Anderson and Shattuck, 2012; Davies et al., 2017). TAPS developed resources, such as lesson plans and pupil work examples, to support the teaching and assessment of science (pstt.org.uk). The principles of formative assessment provide the theoretical basis for guidance to support teacher decision-making and formative action (Davies et al., 2017). When applied to the primary science classroom, these principles emphasize the elicitation of pupil understanding, the responsiveness of teachers to adapt their lessons in response to this information from the pupils, and the active role of pupils in self and peer assessment (Wiliam, 2018). The TAPS resources aim to primarily support teachers with ongoing formative assessment to support learning, but such information can also inform summative assessment judgement as needed in the primary school context.

The TAPS Focused Assessment approach for teaching and assessing scientific inquiry and practical work (called Working Scientifically in the English National Curriculum), proposes that one element of inquiry becomes the focus for teacher attention and any pupil drawing or writing, within the context of a whole inquiry. The TAPS resources provide guidance for suggested activities, but carrying out the activities is not the same as implementing formative assessment. Formative assessment requires action, something needs to be done with the information gained from interactions with pupils, thus teachers need to be open to changing their planned practice and this often requires support that goes beyond downloadable resources.

The three-day Focus4TAPS training program was developed to support the implementation of TAPS and this has been the subject of an Education Endowment

Foundation funded randomized control trial with 2882 pupils across England (Mujtaba et al. 2022). The EEF independent evaluation found a statistically significant impact on pupil learning, equating to two months additional progress for pupils in the intervention group. This paper will take a closer look beyond the headline impact, to consider what changes teachers described making to their practice.

Method

In order to provide a strong, evidence-based argument for schools to join Focus4TAPS program, there needs to be a clear articulation of what difference it makes, hence the research question:

RQ: What do teachers do differently as a result of the Focus4TAPS program?

This paper will draw on the teacher survey data presented in the EEF independent evaluation (Mujtaba et al. 2022), together with qualitative surveys completed by the course attendees during the program, to consider the process of implementation of the training and described impact on practice. The purposive sample required teachers to have attended the Focus4TAPS course. In between training days, the teachers had been asked to carry out some TAPS Focused Assessment activities with their class, which they would then discuss at the next training day. The external EEF survey came at the end of the training. The internal qualitative surveys were completed in the middle of the training program and asked teachers to describe the activities carried out with their class and what they did as a result of such classroom interactions. The teachers were explicitly asked to provide details of any changes to their practice. Such changes can indicate formative use of information: teachers changing their practice in response to information gained from interactions with the children. In line with ethical procedures, all teachers were fully informed regarding the collection, use and storage of their survey answers and were given the opportunity to withdraw their data.

In 2020-21, the three-day Focus4TAPS course ran online as six live online sessions, with follow up activities in between. The external EEF evaluation included an online survey of 111 teachers of Year 5 (age 9-10) in June-July 2021, after all six live sessions had taken place. As part of the ongoing process of development, teachers attending Focus4TAPS training in 2019-20 (n=142) and 2021-22 (n=60) also completed internal surveys asking them to describe their practice. These additional surveys can be used for comparison, especially in terms of the changing landscape across the pandemic school closures. In this brief paper, analysis will focus on the survey questions related to teachers describing changes in their practice, presenting both a raw tally and examples from deductively coded answers (to allow for comparison with a previous study, Earle, 2021).

Results

In both internal and external surveys, for the majority of teachers who answered the question, they described changes to practice (see Table 1). For example, the EEF evaluation found that the majority of Year 5 intervention group teachers who completed the teacher survey response to 'Has Focus4TAPS changed how you teach Working Scientifically?' replied in terms of a positive change, for example, in terms of a more focused/clearer approach, in line with the TAPS 'one focus at a time' approach (Mujtaba et al. 2022: 91). It should be noted that many of the teachers did not choose to answer this free-text question in the online survey (50 answered out of a possible 111). It is likely that this is a mix of survey fatigue (free text takes more effort than ticking a box), but also includes some who are not able to describe a change in practice. Comparison of these EEF evaluation findings with responses to surveys in other years, provides further evidence of teachers describing changes in their practice. When teachers in the 2019-20 and 2021-22 groups were asked about use of a TAPS activity, the majority described a change in Working Scientifically practice (see Table 1).

Table 1. Teacher survey question regarding change in practice

Year and mode of training	Number of teacher responses	Number describing change in Working Scientifically	Number describing no change
2019-20 – face-to-face	142	130	12
2020-21 – online (EEF)	50 (61 non-answers)	47	3
2021-22 - blended	60	54	6

Beyond descriptive statistics, the internal and external evaluations gathered information about what kinds of changes teachers made in their practice. For example, in the EEF evaluation, one teacher who was interviewed, described how the new teaching focus helped to open out the investigation for the children:

> I suppose it made me a little bit more [focused]. I'm focused and how I was approaching things in particular, so I guess certain terminology was being a lot more focused … in terms of their own independent investigation, so come and letting them go free. You could actually see how they were working things out for themselves a lot better. So yeah, definitely it was helpful … Yeah much, much better [at being confident in teaching science and assessing children Working Scientifically] because I know what I'm looking for now, it just it just [sic] makes it makes sense (Mujtaba et al. 2022: 91).

The ’letting them go’ represents a move away from recipe-driven inquiry, where the aim would be to follow the instructions, to more open inquiry, where the children are able to make more decisions to answer their question. By selecting a focus for the lesson, the teacher could give the children choice in the other parts. For example, if the focus was on drawing conclusions, then writing a provided method and results in a strict format would not be necessary, releasing time for more discussion about the findings.

In the internal evaluation survey, during Day 2, the teachers were asked: ‘What did you do next (in response to this lesson)?’ An identical question was asked of the 2019-20 groups, so a similar coding was used to categorise the spread of formative action (Earle, 2021). Some brief examples are presented in Table 2.

Table 2. Examples of formative action from 2021-22 groups (N=60)

Timing of described formative action	Type of formative action described	Example 1	Example 2
1. As part of the same lesson or activity	a. More time on pupil recording of investigation e.g. draw diagram, complete table/ graph, write prediction/ conclusion	Supported children drawing tables and looking for rogue results. T170	Discussion of the best biscuit. Write up. Whole class attempt with the TAPS recording system. T224
	b. More time for discussion with pupils of results, conclusions, evaluations or reflections	Reviewed the string telephones. Children determined the best design and justified why this was the case - Material used or length of string etc. T166	Children showed work as a gallery so they could gain ideas from each other. Reviewed children’s posters to identify misconceptions, then addressed these in lessons. T229
2. Activity follow up (extended or next lesson on same topic)	a. New focus on vocabulary	Use vocabulary discussed to pick up in the next topic. So build on this learning. T168	During the next lesson and throughout the rest of the unit key vocab was recapped and links were made to the concepts covered. T228
	b. New focus on concepts	Further study of animals skeletons, noticing similarities and differences. T190	Review classes of animals, create classification key to ID types of animals. T234

3. Adapt future teaching	a. Explicit teaching or practice of science skill	It will enable me to structure the coming lessons on developing questioning skills and knowing who I need to push further if they are already fairly confident. T178	Continue to use the planning boards to build confidence in use. Build in regular time to ask Qs and plan enquiry. T245
	b. Modelling, scaffolding or use of examples	Looked at good examples and modelled how to improve predictions. T165	Rather than getting the children to complete independently, I fully modelled how to draw the table and the graph and we talked through step by step. T253

Table 2 demonstrates that some teachers might make an immediate (within the lesson) change to planned practice, whilst others may use the information gathered from the children to adapt future plans. The exact timings would be dependent on a wide range of factors, such as whether the lesson fell at the beginning or end of a topic. The timing and nature of a formative action would be a professional decision regarding the adaptation of teaching to match the learning needs of the children. It is making any change, in response to ongoing observation and assessment, that makes it formative use. Usual practice in a primary school in England would be to move on to the next pre-planned lesson, which is what has sometimes been noted if teachers download the TAPS resources without attending the Focus4TAPS programme. The time to think and reflect upon how to tweak their practice, afforded by a professional development program, helps teachers to build formative actions into practice.

The examples described in Table 2, describe whole class adaptations, rather than adjustments for individual children. This may be because, for time-poor primary science, a 'best-fit' action for the whole class is the most manageable outcome. There may have been individual adaptations or scaffolds provided for children with additional needs, but this was not explicitly asked about in the survey, so could be an area for future research.

Discussion and Conclusions

The TAPS Focused Assessment approach has been presented across England as a manageable way for primary school teaching and assessment of science inquiry. The internal and external surveys both provide evidence of use of the approach and described impact on practice. Nevertheless, a sustained professional development program was necessary to support such teacher understanding of primary science and practical inquiry, together with formative use of assessment information gained from interaction with the children. The lack of time, money and supply cover available means that even with such robust evidence of impact,

schools in England are still struggling to take part in the program. In addition, the retention crisis means that even where schools have joined the program, they may need to re-visit the training due to staff changes.

A further concern is the way such a program could become transmissive rather than malleable (Kennedy, 2014), with teachers being the receivers of information, rather than agentic partners in their own professional development. The addition of collaborative elements from the original Design-Based Research project are possible in the program, with opportunities for new examples to be developed, but whether teachers have time for this, and whether this is enough to develop teacher agency, remains to be seen. Randomized control trials need a 'product' to test, so further research is needed to see whether that 'product' can be malleable enough to empower teachers, as well as enable the implementation of a new approach to support the teaching of science inquiry.

References

Anderson, T., & Shattuck, J. (2012). Design-based research: a decade of progress in education research? *Educational Researcher*, 41(1), 16-25.

Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B. & Wong, B. (2010). '"Doing" science versus "being" a scientist: Examining 10/11-year-old schoolchildren's constructions of science through the lens of identity', *Science Education*, 94, (4), 617–639.

Association for Science Education (ASE) (2020). *Best Practice Guidance: Guidance on Timetabling Science within the Primary Curriculum.* https://www.ase.org.uk/system/files/BPG%20Timetabling%20Primary%20Science.pdf

Bianchi, L, Whittaker, C.& Poole, A. (2021). *The 10 key issues with children's learning in primary science in England.* The University of Manchester and The Ogden Trust.

Davies, D., Earle, S., McMahon, K. Howe, A. & Collier, C. (2017). Development and exemplification of a model for Teacher Assessment in Primary Science, *International Journal of Science Education*, 39:14, 1869-1890.

Earle, S. (2021). Formative Decision-Making in Response to Primary Science Classroom Assessment: What to do Next? *Frontier Education*, 5:584200.

Kennedy, A. (2014) Understanding continuing professional development: the need for theory to impact on policy and practice, *Professional Development in Education*, 40:5, 688-69.

Nag Chowdhuri, M., King, H. & Archer, L. (2021). *The Primary Science Capital Teaching Approach: Teacher handbook.* London: UCL Institute of Education. https://discovery.ucl.ac.uk/id/eprint/10136335/

Mujtaba, T., Sheldrake, R., Hodgen, J. & Reiss, M. (2022). *Focus for Teacher Assessment of Primary Science (Focus4TAPS) Evaluation Report.* London: Education Endowment Foundation. https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/focus4taps

TAPS website: https://pstt.org.uk/unique-resources/taps/

Wellcome (2017). *State of the Nation report of UK Primary Science Education.* London: The Wellcome Trust.

Wiliam, D. (2018). *Embedded Formative Assessment* (2nd ed). Bloomington: Solution Tree Press.

Plant and Animal Photographs in the Science and Environment Textbooks of the Greek Primary School

Georgios Ampatzidis[1], Alexandros Amprazis[2] and Penelope Papadopoulou[2]

[1]University of Thessaly, Greece

[2]University of Western Macedonia, Greece

The connection between humans and nature has been a recurring topic of discussion, occasionally identified as a primary factor contributing to environmental degradation. Within this context, there has been a particular emphasis on examining people's engagement with plants, with a specific focus on a phenomenon referred to as 'plant blindness' or 'lack of plant awareness'. Even though a diminished interest for plants at first thought may not seem so important, plant blindness can have some major implications such as a possible link to sustainable development. Disregarding the significance of plants can act as a counterforce to environmental equilibrium and actively impede progress toward attaining the majority of Sustainable Development Goals. Research shows that plants do not seem to have a strong presence in students' everyday life and a general zoo chauvinism seems to be established in educational institutions to the detriment of botanical education. Aim of this research was to examine the presence of plants in the photographs included in the Greek primary school science and environment textbooks. We investigated six textbooks and we identified 948 photographs in total which we coded in mutually exclusive categories. Our results revealed that the predominant subjects in the photographs of the textbooks we examined were humans. Moreover, among humans, plants and animals, plants were the least commonly featured. We discuss our results in the light of the relevant literature.

Keywords: Plant blindness, science and environment textbooks, primary education

Introduction

Humans' connectedness to nature has always been a matter of discussion and sometimes has been considered as one of the main environmental degradation causes. Within this framework, there has been a focus on people's interaction with plants and especially on a phenomenon called 'plant blindness'. This phenomenon, recently renamed as 'plant awareness disparity' or 'lack of plant awareness' (Pany et al., 2022), refers to people's tendency to neglect

and underestimate plants. To be more specific, there seem to be deficiencies in humans' attention, attitude, knowledge, and relative interest towards plant organisms (Parsley, 2020). Even though a diminished interest for plants at first thought may not seem so important, plant blindness can have some major implications such as a possible link to sustainable development (Thomas et al., 2021). Ignoring plants can be an opposing factor to environmental balance and directly hinder the achievement of the vast majority of Sustainable Development Goals (Amprazis & Papadopoulou, 2020). Identifying and appreciating all organisms upon Earth is fundamental to thoroughly comprehend biodiversity and act accordingly.

Regarding plant blindness' causes, a variety of reasons has been reported in the literature throughout the years (Stagg & Dillon, 2022). Apparently, humans' resemblance to other mammals makes it easier for people to relate or feel more familiar with animals than plants. Humans' brain selectivity regarding the incoming info is also listed among the phenomenon's causes. Plants, as a common part of the visual background and without visible movement, are not usually among the information that the brain chooses to process (Achurra, 2022). Additionally, some researchers have suggested that the educational systems worldwide are in a sense more 'animal-oriented' and less 'plant-friendly'. Plants do not seem to have a strong presence in students' everyday life and a general 'zoo chauvinism' seems to be established in educational institutions to the detriment of botanical education (Stroud et al., 2022). As this deficient educational approach that enhances the plant blindness phenomenon has been identified, it is crucial to investigate it further and reverse the situation. Analyzing the curriculum and textbooks is a critical step in forming comprehensive conclusions regarding this matter within any educational setting.

In regard to everything that has been mentioned above, aim of this research is to examine plants presence in Greek primary school science and environmental textbooks. More specifically, we are interested in the presence of plants in the photographs included in the aforementioned textbooks. So, the research question of this study is whether there are different numbers of photographs of plants and animals in the Greek primary school science and environment textbooks.

Methods

In the framework of this study we investigated the four textbooks used for the subject "Study of the Environment" (Grade-1 to Grade-4) and the two textbooks used for the subject "Natural Sciences" (Grade-5 and Grade-6) during the school year 2022-23. More specifically the textbooks we researched are the following:

1. Study of the Environment for the 1st Grade of Primary School (Plakitsi et al., 2015)

2. Study of the Environment for the 2nd Grade of Primary School (Dimopoulou et al., 2015)
3. Study of the Environment for the 3rd Grade of Primary School (Kokkotas et al., 2015)
4. Study of the Environment for the 4th Grade of Primary School (Kokkotas et al., 2016)
5. Natural Sciences for the 5th grade of Primary School (Apostolakis et al., 2015a)
6. Natural Sciences for the 6th grade of Primary School (Apostolakis et al., 2015b)

We note that 'Study of the Environment' and 'Natural Sciences' are two subjects that focus on science and the environment at the primary school level; the former is offered throughout the first four grades while the latter is offered at the 5th and 6th grade.

Each photograph in the textbooks that featured plants or animals served as the study's unit of analysis. Diagrams and drawings were excluded from the study's analysis since they are abstract representations that are often used to illustrate concepts or processes. In cases where a non-living object served as the main subject of the photograph, we also disregarded unintentional appearances of plants or animals (e.g. some plants in front of a building). Despite of belonging to Kingdom Animalia, humans were considered as a separate unit analysis and required additional coding. That methodological approach was chosen according to the literature regarding plant blindness, as humans were separated by researchers from the other animals to create distances between them and animals/plants. That separation was necessary in order to study precisely the differentiation between these two distances and draw conclusions regarding a possible bias.

We identified 948 photographs in total which we coded in mutually exclusive categories formed drawing on a scheme proposed by Link-Pérez et al. (2010). More specifically, the categories concerning the photographs were the following: animal subject, human subject, plant subject, landscape subject, and dual subject (human and plant, animal and plant, human and animal) (Table 1). The first and third author coded independently the photographs included in the first two textbooks (Grade 1 and Grade 2) and the rater agreement was about 91% for textbook-1 and about 98% for textbook-2. The rest of the analysis was carried out by the first author.

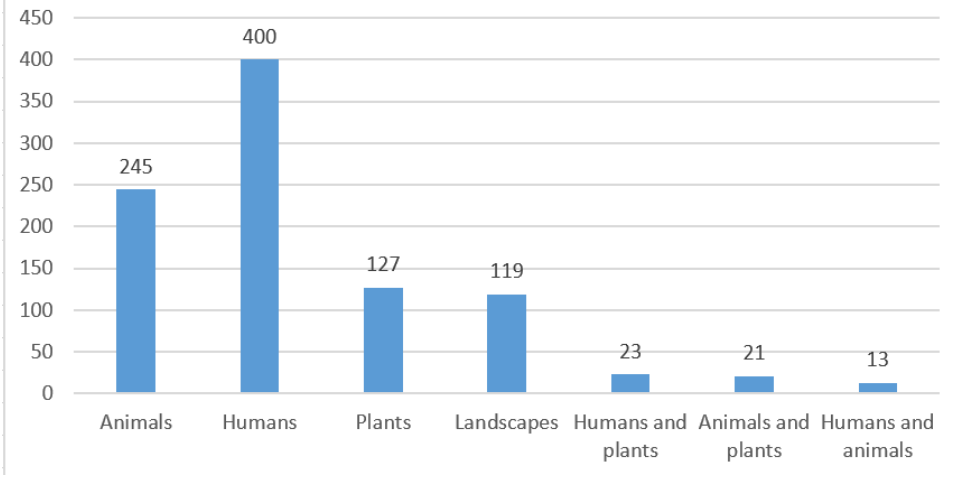
Table 1. The coding scheme.

Subject of the photograph	Examples
Animals	The main subject of the photograph is one or more animals
Humans	The main subject of the photograph is one or more humans
Plants	The main subject of the photograph is one or more plants
Landscapes	The main subject of the photograph is an abstract landscape
Humans and plants	The main subject of the photograph is one or more humans along with one or more plants
Animals and plants	The main subject of the photograph is one or more animals along with one or more plants
Humans and animals	The main subject of the photograph is one or more humans along with one or more animals

Results

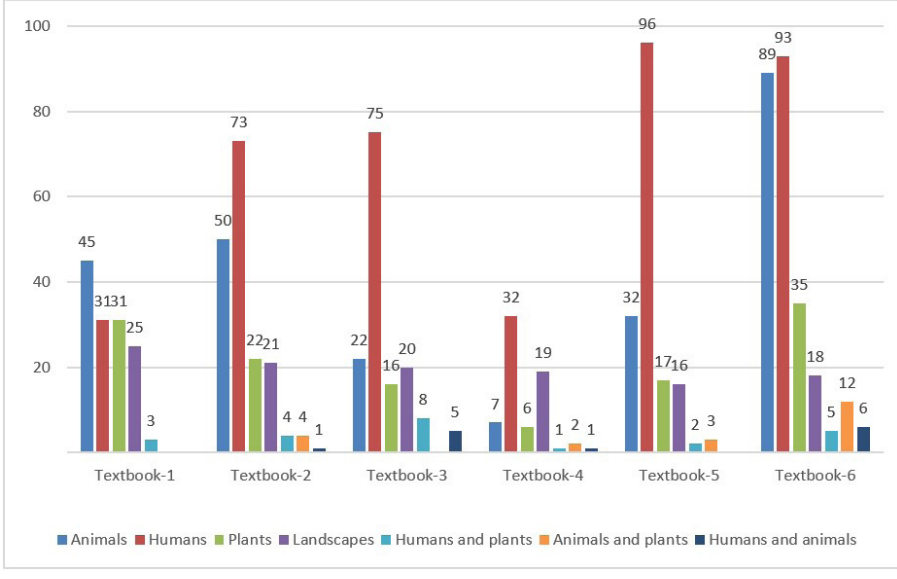
Our results showed that humans were the most frequent subjects of the photographs of the textbooks we investigated. Among humans, plants and animals, plants were the least popular. In the 948 photographs that were coded in all six textbooks, 245 had animals as their main subject, 400 had humans as their main subject, 127 had plants as their main subject, 119 had a landscape as their main subject, and 57 had a dual main subject (Figure 1).

Figure 1. Frequencies of categories of photographs' subjects per textbook.



Focusing in different textbooks, we noticed that photographs with humans as their main subject are more in five out of six textbooks. Moreover, in all six textbooks photographs with plants as their main subject are less than the ones with animals as their main subject (Figure 2).

Figure 2. Frequencies of categories of photographs' subjects per textbook.



Conclusions

According to the results, Greek primary education textbooks are not all-encompassing and balanced regarding plants' references in them. Comparing to animals and humans as a distinct category of the Kingdom Animalia, the vegetal world seems to fall short. Even if this was not designed by purpose, an issue is identified and should be addressed. It is highly probable that this bias was created due to the low levels of plant awareness that unconsciously seems to characterize our social structure and the educational community (Bobo-Pinilla et al., 2023). Redesigning and enriching Greek textbooks in favor of plants, can be a counter action to this zoo chauvinism and to the plant blindness phenomenon. Education can be both a cause and a solution to humans' restricted interest in plants. Changing the reduced emphasis on plant life in the Greek educational system and worldwide can be a solid starting point for reconverting people's attention and appreciation for plants. As already mentioned, restricting plant blindness is crucial for understanding the significant role plants play for maintaining life on planet Earth and their undoubtable contribution to human welfare (Lawrence & Calvo, 2023).

If one of the major aims of education is to enhance students' environmental citizenship through reestablishing their connectedness to nature (Restall & Conrad, 2015), specific requirements are necessary. Diminishing phenomena like plant blindness, uprooting notions of popular and unpopular animals, embracing and comprehending the importance of biodiversity, are all goals that cannot be achieved through occasional educational interventions. In a modern school that

can support such a transformative learning, environmental conservation and sustainability are not just teaching subjects, but principles upon which the whole institutions' operation is organized. In that context, curriculum and textbooks are very crucial, as they constantly form students' area of interest and background knowledge. Therefore, it is totally recommended for policy makers to re-evaluate textbooks and assure that they conform to a holistic approach regarding the natural environment and the sustainable education conceptual framework.

References

Achurra, A. (2022). Plant blindness: A focus on its biological basis. *Frontiers in Education*, *7*. https://doi.org/10.3389/feduc.2022.963448

Amprazis, A., & Papadopoulou, P. (2020). Plant blindness: A faddish research interest or a substantive impediment to achieve sustainable development goals? *Environmental Education Research*, *26*(8), 1065–1087. https://doi.org/10.1080/13504622.2020.1768225

Apostolakis, E., Panagopoulou, E., Savvas, S., Tsagliotis, N., Pantazis, G., Sotiriou, S., Tolias, V., Tsagkogeorga, A., & Kalkanis, G. (2015a). *Natural Sciences for the 5th grade of Primary School*. CTI Diophantus.

Apostolakis, E., Panagopoulou, E., Savvas, S., Tsagliotis, N., Pantazis, G., Sotiriou, S., Tolias, V., Tsagkogeorga, A., & Kalkanis, G. (2015b). *Natural Sciences for the 6th grade of Primary School*. CTI Diophantus.

Bobo-Pinilla, J., Marcos-Walias, J., Delgado Iglesias, J., & Reinoso Tapia, R. (2023). Overcoming plant blindness: Are the future teachers ready? *Journal of Biological Education*, *0*(0), 1–15. https://doi.org/10.1080/00219266.2023.2255197

Dimopoulou, M., Zompolas, T., Bampila, E., Skanavi, K., Frantzi, A., & Chatzimichail, M. (2015). *Study of the Environment for the 2nd grade of Primary School*. CTI Diophantus.

Kokkotas, P., Alexopoulos, D., Malamitsa, A., Mantas, G., Palamara, M., & Panagiotaki, P. (2015). *Study of the Environment for the 3rd grade of Primary School*. CTI Diophantus.

Kokkotas, P., Alexopoulos, D., Malamitsa, A., Mantas, G., Palamara, M., Panagiotaki, P., & Piliouras, P. (2016). *Study of the Environment for the 4th grade of Primary School*. CTI Diophantus.

Lawrence, N., & Calvo, P. (2023). Learning to See 'Green' in an Ecological Crisis. In L. Weir (Ed.), *Philosophy as Practice in the Ecological Emergency: An Exploration of Urgent Matters* (pp. 167–183). Springer International Publishing. https://doi.org/10.1007/978-3-030-94391-2_9

Pany, P., Meier, F. D., Dünser, B., Yanagida, T., Kiehn, M., & Möller, A. (2022). Measuring Students' Plant Awareness: A Prerequisite for Effective Botany Education. *Journal of Biological Education*. https://doi.org/10.1080/00219266.2022.2159491

Parsley, K. M. (2020). Plant awareness disparity: A case for renaming plant blindness. *PLANTS, PEOPLE, PLANET*, *2*(6), 598–601. https://doi.org/10.1002/

ppp3.10153

Plakitsi, A., Kontogianni, A., Spyratou, E., Manoli, V., Balla, E., & Starakis, I. (2015). *Study of the Environment for the 1st grade of Primary School*. CTI Diophantus.

Restall, B., & Conrad, E. (2015). A literature review of connectedness to nature and its potential for environmental management. *Journal of Environmental Management*, *159*, 264–278. https://doi.org/10.1016/j.jenvman.2015.05.022

Stagg, B. C., & Dillon, J. (2022). Plant awareness is linked to plant relevance: A review of educational and ethnobiological literature (1998–2020). *PLANTS, PEOPLE, PLANET*, *4*(6), 579–592. https://doi.org/10.1002/ppp3.10323

Stroud, S., Fennell, M., Mitchley, J., Lydon, S., Peacock, J., & Bacon, K. L. (2022). The botanical education extinction and the fall of plant awareness. *Ecology and Evolution*, *12*(7), e9019. https://doi.org/10.1002/ece3.9019

Thomas, H., Ougham, H., & Sanders, D. (2021). Plant blindness and sustainability. *International Journal of Sustainability in Higher Education*, *23*(1), 41–57. https://doi.org/10.1108/IJSHE-09-2020-0335

"It's Only a Drawing!" - Learning to Model and About Models in Primary School

Francesco Cuomo[1], Emilio Balzano[2], Anna Merinio[3], Marco Serpico[3]

[1]Alice Salomon University of Applied Sciences, Germany

[2]Università degli Studi di Napoli Federico II, Napoli, Italy

[3]Università degli Studi Suor Orsola Benincasa, Napoli, Italy

Modeling is considered a key scientific practice to be addressed in science education at all grades. Current performance expectations propose that children in the early grades can learn to develop simple evidence-based models to represent objects or describe phenomena, while higher level competences related to more abstract models and their use to describe relationships among variables between systems and their components are supposed to be accessible in high school. One of the three components of modeling competence, as identified in recent literature, knowledge about the use of models, their strengths and limitations, defined as meta-modeling competence, is not addressed in the performance expectations we take as reference (the Next Generation Science Standards). In a design study involving ca. 100 children over two years, we investigate the pedagogical settings that promote the development of modeling competence and the related emerging students' levels of competence in the early grades of primary school. Our findings suggest that children in second and third grade are able, in appropriate pedagogical settings, to learn to model and about models to a degree of abstraction and complexity that exceeds the current performance expectations, including developing meta-modeling knowledge.

Keywords: Scientific practices, modeling, primary school.

Introduction

Modeling can be considered a central, defining practice in science (Passmore, Stewart, & Cartier, 2009). As such, it is of paramount relevance also in science education at all grades. The National Research Council (2012) lists modeling as one of the eight most relevant scientific practices that identify one of three dimensions in science education and should be addressed in order to "cultivate students' scientific habits of mind" (p. 41). In a recent development in the international research discourse on science education, scientific practices such as argumentation and modeling – formerly "underemphasized in the context of science education" (NRC, 2012, p. 44) – are gaining a growing relevance

(Kenyon et al., 2008; Lehrer & Schauble, 2006, 2015; Schwarz et al., 2009). This is seen as an improvement with regards to a wide spread tendency to reduce inquiry to experimental investigation and "a single set of procedures, such as identifying and controlling variables, classifying entities, and identifying sources of error" (NRC, 2012, p. 43). Modeling competence is defined as a system of the knowledge, skills, and abilities necessary to engage in the process of developing and using models for reasoning in science (Upmeier zu Belzen et al., 2019). Recent literature (Chiu & Lin, 2019; Nicolaou & Constantinou, 2014; Nielsen & Nielsen, 2021) has further specified modeling competence along three components related to modeling practice, modeling product and metamodeling knowledge, the latter corresponding to knowledge about "how models are used, why they are used, and what their strengths and limitations are" (Schwarz et al., 2009). The performance expectations for grades 2-3 of primary school as identified in the Next Generation Science Standards (NGSS Lead States, 2013) involve developing simple models based on evidence to represent a proposed object or tool (p. 21) or describe phenomena (p. 27). It is not before high school that students are expected to be able to develop and use evidence-based models to "illustrate the relationships between systems or between components of a system" (p. 92). The Standards make no reference to performance expectations related to meta-modeling knowledge. This article aims at investigating teaching-learning designs that promote the development of modeling competence in the early grades of primary school, the modeling practices that children use when exposed ton these designs and their emerging levels of competence.

Research Setting and Methods

Our findings emerge from an action-research programme implemented in a primary school in a suburban area of a large city in Italy (ca. 1 million inhabitants). Four classes (ca. 100 children) were involved in the programme over a period of five years. This corresponds to the complete primary school cycle in the Italian education system. In each school year, seven interventions were carried out (ca. 1.5 h each) by the researchers with the participation of the class teachers. The interventions were all centred on authentic phenomenological contexts and we (and the teachers) restrained from direct instruction of models. Our methodological reference is design research, an approach that aims at developing empirically grounded theories through a combined study of both the process of learning and the means that support that process (e.g., Cobb et al. 2003; diSessa & Cobb, 2004). A design research approach implied that we designed and implemented interventions in authentic classroom contexts, whereas the regular presence of the researchers in the classroom over an extended time-frame allowed for a seamless integration of the intervention in the students' everyday school experience. Children were involved in hands-on inquiries in different phenomenological domains and encouraged to produce enactive,

iconic and symbolic representations of their ideas in interaction with their peers, the researchers and the teachers. During their investigations, children worked both individually and in small groups. Each activity was concluded with ample time devoted to a whole class discussion in which the children shared their ideas and representations in a process of collective sensemaking (Odden & Russ, 2019) and co-construction based on a sustained shared thinking approach (Siraj-Blatchford et al., 2002). The iterative nature of the design approach involved a process of cyclic design, evaluation and revision in which interventions were planned according to provisional assessment of learning outcomes. This implies that learning and design can be studied in their mutual interdependence. Data collection included video- and audio-recordings of the interventions and field notes. Also, the inscriptions produced by the students were partly photographed and partly collected and scanned. Selected data (focussing on activities related to modeling) were analyzed qualitatively according to a content analysis approach (Hsieh & Shannon, 2005; Lichtman, 2013; Mayring, 2020).

Findings

This paper focusses on findings related to the activities carried out in the second and third grade. The investigation contexts ranged from pouring water between containers of different size and form, the behaviour of objects when immerged in water (floating and sinking), exploring different materials (in terms of volume, weight and the relationship between them) and the graphic representation of the different states of an inflating and deflating balloon over time. In order to illustrate the observed learning processes in their complexity we choose to use the limited space available herein to present two exemplary sequences extensively.

Sequence 1. Understanding the conical container

In the second grade, the activities started in a learning environment prepared with a large number of containers of different size and form (including variable height and section), some of which were filled with water. After a period of free exploration, the children are given the task of comparing the amount of water in the containers (the capacity or "size" of the containers). As this is designed as a whole class activity and a consensus has to be found with regards to how to address the task, the children resort to a variety of arguments to support their hypotheses. Some children point out that some containers are higher and can therefore contain more water. "But it can also be very short and very wide," another child points out, to which another adds: "some of them contain the same amount: [mimicking with his hands] that one is larger on the bottom, then widens and then shrinks again." A few minutes into their negotiations, the children realize that a more precise strategy is necessary to address the task more effectively. Also the strategy to be adopted is the subject of negotiation. A. proposes to "empty the one we think is the largest and pour [the content of] all the other ones in it."

A similar strategy emerges in another class, where F. proposes to "take one jug and pour the water in a container and then we measure the highest water level [in the container] for all these jugs but we take a marker pen to see when we get highest." Both these strategies reflect modeling practices implemented in order to address the question "how much" in an objective fashion that is typical of science. Children have identified and control the water volume variable using a consistent criterion for making comparisons between different phenomenological situations, providing assessments of "more," "less" or "equal." Both A. and F. have done so by developing a sort of measurement theory based on relationships between different units (the containers). This strategy finds consensus in both classes and the children proceed to measure many different quantities of water. By repeatedly pouring water from the small containers (100 ml) and marking the successive levels of the liquid in the larger container they obtain a graduated cylindrical container. In a later sequence, the researchers introduce a graduated container of conical shape (the horizontal section increasing with the height, as depicted in Fig. 2). The children immediately notice that in this case, contrary to their graduate cylindrical container and other commercial graduated containers observed earlier, the grades on this conical container are not regularly spaced. After having verified that the marks are correct by pouring water from the 100ml container, the children are puzzled by the different grading structure. In a whole class discussion, they are asked to draw the container on the blackboard to see if this helps understanding the puzzling phenomenon. They start using the drawing in order to develop arguments about the decrease of the distance between the marks with the height. One child starts manipulating the drawing in an attempt to show her classmates that the different areas between markings are equal. To do so she uses a kind of reasoning that capitalizes on the intuitive observation verbalized in the first, explorative phase of the activity that containers of different height might contain the same amount of water because of the different section. This confirmed our expectation that the design of the learning activities should involve starting by engaging the children in simpler cognitive tasks involving the comparison of two artifacts each with fixed width before introducing the artifact of variable width. She proposes to "cut" the lower section (between 0 and 100ml) in two halves horizontally and then "move" them both next to the upper section (between 300 and 400ml) of the container as shown in Fig. 1 (Fig. 2 provides a diagrammatic representation of the operations carried out by the student on the blackboard.) Some of the children now manifest their agreement with L.'s theory, while others seem to be still confused. E.: "If we have to do this [move one half of the lower section], we cannot do it because it is not so that water can fly!" To which F. responds: "But this is a drawing that we are trying to let us understand this [by mimicking the shifting of the stripes done previously] because this one can be made to understand. This is a mock drawing! It's a figure of speech!" L: "It's a way of speaking by drawing. Do you understand?"

Researcher: “that is, E. is saying «but you are talking about a series of maneuvers with water which are very strange: you want to cut it, you want to make it fly», right, E.? Is this what you were saying? «We can’t do that!» But, if I get it right, L. and F., you don’t really want to do these operations with water…” L. “ooooh! it’s only a drawing, only a drawing!”

Figure 1. L. illustrates her theory for the equality of amounts of water in two sections of the graduated container of conical shape by manipulating the drawing.

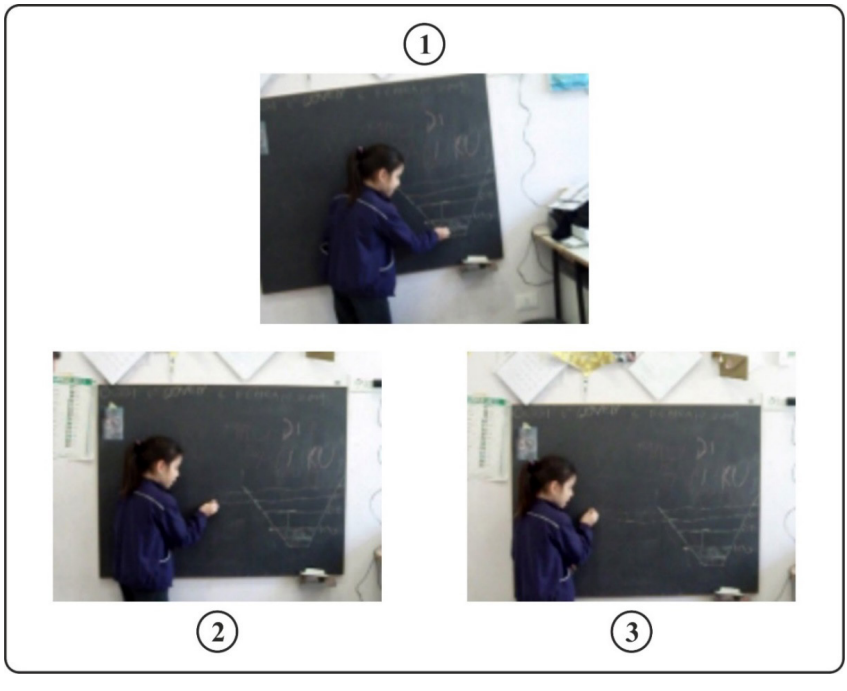
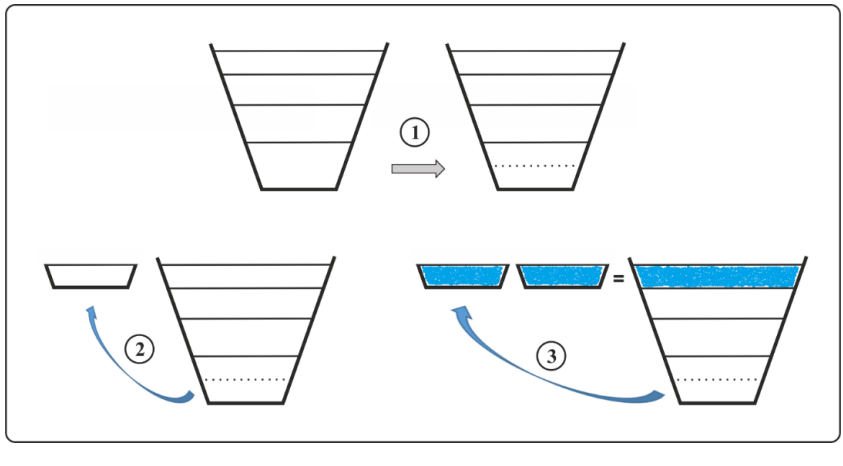


Figure 2. A diagrammatic representation of the manipulations carried out by L. on the drawing.



In this sequence, the children have designed and manipulated a modeling product in order to make sense of a puzzling phenomenon by developing a theory that the amounts of water in the two sections (real world) are the same because the areas of the two sections (model world) are the same. To do this they have carried out a series of operations on the modeling product that reflect operations on the mental models of highly abstract nature. At the same time, they have stated the strengths of the modeling product (it can "let us understand"; "it's a way of speaking by drawing") and its limitations ("it's a way of speech"; "it's only a drawing!"), therefore clearly demonstrating meta-modeling knowledge that was not explicitly taught to them.

Sequence 2. The story of a balloon

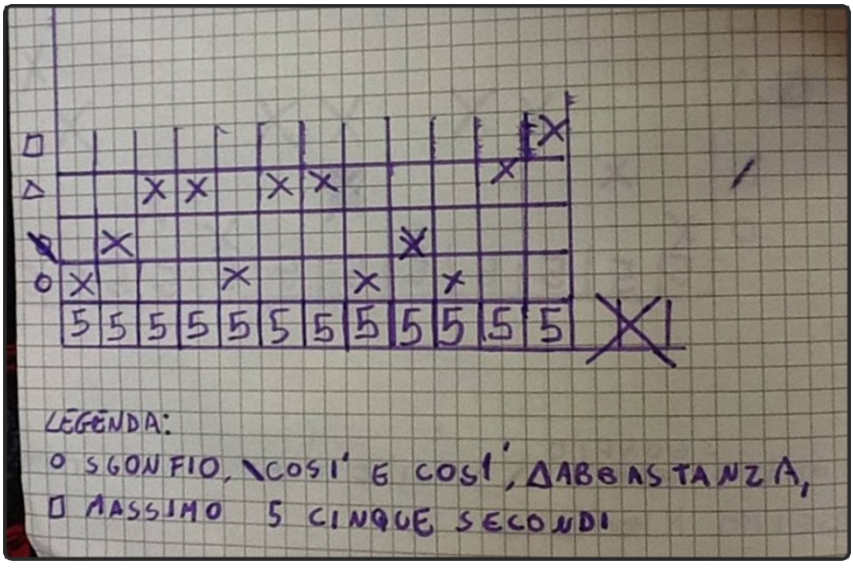
In third grade, the children were involved in developing some form of representation for a physical transformation. This implies the search for forms that can account for two variables since a transformation in the natural world is a change in one aspect of a phenomenon with respect to time. The design of this activity aimed at fostering the development of more abstract graphical representations resembling the Cartesian coordinate system without resorting to direct instruction. As of the phenomenological domain, we chose a context that would entail a dimension of play. Children were invited work in pairs and observe the transformations of a balloon as it is inflated and deflated. While one would freely play with the balloon, inflating and deflating it by different amounts of air in different time intervals, the other has the task of designing a table or some kind of inscription in which the transformations of the balloon can be "recorded." Children usually learn to use different kinds of tables in the first grades of primary school. One type of table is especially well known to children at this age: the calendar, which also usually hangs in every classroom in Italian primary schools. In this particular school, a specific version of the calendar is used to document the attendance of each child. Every day, the children present in class attach their specific sticker – which they could choose at the beginning of the year – to the cell corresponding to the current day (see Fig. 3 below).

Figure 3. The table used in the school setting of this study to mark children's attendance. The children names are listed in the rows, while the columns represent the days.



Before the work in pairs begins, we moderate a whole class discussion aimed at deciding how to keep track of the flowing time. In all classes, the children decide to do this by counting also soon come to the conclusion that this has to be done at a regular rhythm. In some classes, children decide to sing the numbers, in others they clap their hands at regular intervals while counting. The ideas related to how to represent the different sizes of the balloon are more varied. Some children want to use an iconic representation that visually reproduces the balloon, others more abstract symbols, the meaning of which they would explain in what we would call a legend complementing the table. In every class, the children decide to set a fixed number of counts (five in some classes, ten in others) to which each cell in the table should correspond. They call each count a second. In analogy to the days reported in the attendance table they represent this variable horizontally. On the vertical axis, they report the different sizes of the balloon. Interestingly, all children order the vertical axis with increasing heights corresponding to increasing values of the size of the balloon, even if this feature of the representation could not be borrowed from the attendance table. In the example documented in Fig. 4, the children used abstract symbols to represent four different sizes of the balloon. A circle is referred to the completely deflated balloon, a circle with a slash represents "so so," a triangle "enough" and a square "maximum." They then use an X to indicate the size the ballon in each time interval.

Figure 4. An example of the inscriptions designed by the children to represent the changes in size of a balloon over time.



When the children have completed their first task, we switch back to the whole class setting to reason on meaning of their inscriptions and their effectiveness in describing the phenomenon. To do so we select some tables that seem particularly promising for a fruitful discussion and we ask the children to reproduce them on the blackboard. For the whole group, we set the task of finding out how to "read" this inscription. The following example refers to the representation produced by L. and F. and shown in Figure 5 below.

L.: *We put the X where the… size… of the quantity arrived.*

Teacher*: Could you explain a bit more what you mean with that?*

L.*: It is the…*

F.*: It is the measure of the balloon.*

Researcher*: What measure?*

Various children*: of the volume*[1] *… L. said of the size.*

L.*: "We put the X after the first ten seconds in which the balloon had inflated, then we did another ten seconds but we did not deflate the balloon and then we counted another ten seconds… From the volume of the balloon we added another volume."*

Researcher*: that is, you inflated more…*

1 The children in our sample had encountered the concept (and the term) of volume in the activities conducted in grade 2, when they used graduated containers and scales to measure volume and weight of difference materials.

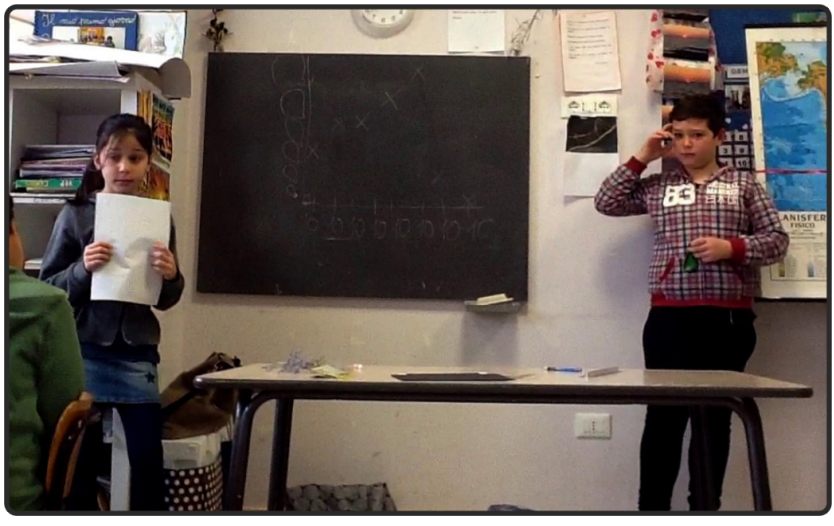
L.*: Then in the third [interval] we left it [the volume] the same.*

F.*: Then in the fourth we found the superior volume and in the fifth as well but then we deflated it and we have the X there [pointing at the X in the sixth column]*

Researcher*: Was it not completely deflated?*

L.*: No, it arrived here [pointing at the X in the sixth column] and then it became deflated in the last one.*

Figure 5. Children describe their representation of the transformations of the balloon at the blackboard in a whole class discussion.



The diagram realised by the two children can be seen as a prototype of the Cartesian volume-time graph: on the y-axis the variable volume is represented by drawings in a coherently increasing order, on the x-axis the equal and regular time intervals are represented as a series of 'tens', with L. specifying that each successive ten is 'another ten', following to the previous ones. The children are appropriating this new form of representation, which is a sort of discretized volume vs. time graph. The next cognitive step is to move gradually from the discrete to the continuous. Having chosen one of the children's charts and reproduced it on the blackboard, we draw a line passing through the crosses (see Fig. 6) and invite the children to reason on the possible meanings of this new element. In particular, we invite them to say whether it is an enrichment of our representation of what the observed phenomenon, if it tells us something more about it. Some of the children follow the line drawn on the blackboard with their finger and at the same time verbalize what this representation suggests. C. says she is going to "tell the story of the balloon" and she does so by starting from the highest point on the graph (the volume at 50 seconds, Fig. 6) and moving along the curve to the left: "here the balloon was really inflated, here it was quite

inflated, here it was so so, here it was a little bit like that" [in fact these two crosses are at same height]. We ask her to start again, addressing the questions "where do we start?" and "how do we follow the line?". She shows the highest cross of all and says: "from here" then runs her finger along the line from right to left. As she observes the drawing more carefully, she notices, however, two more crosses to the right of the highest point and adds: "ah, here's another one... and so instead of starting from here we should start from here" [from the far right of the graph]. At this point, some children begin to express dissent at her explanation. One of them, A., joins her at the blackboard and says that the drawing should be handled like a book: the story should be read from the first to the last page. To which he adds: "you have to start from the smallest number, because the smallest number is what happened before."

Figure 6. Children reason about the meaning of a continuous line connecting their "experimental" points.



The children carry out their argumentations further, as the discussion is moderated by a researcher:

A.: *C., you can't do that because otherwise we would have done it the other way round, from 70 to 10.*

Researcher*: But 10, 20, 30, 70… what are they?*

C.*: Seconds.*

Researcher*: Ok, so it's the time ... so what does A. mean with "the other way around?"*

A.*: I mean that from inflated… that is…*

A. stares at the graph and seems unable to say anything more. Interestingly, it is C. herself who helps him explain further.

C.*: You mean that if I wanted to start from here* [indicating the point where the balloon's volume is at its maximum]*, 70 should be here* [indicating the first "box" of the table on the left of the graph, corresponding to the first 10 seconds].

F.*: But no, you must always start from the smallest number.*

A.*: Yes, it is like a book … you can start from the first page you cannot read from, from…* [pointing at 70]

Researcher*: And the beginning of this story is what?*

A.*: From here* [pointing at 10]

The new element introduced in our educational design leads the children to reason on both the syntactic and the semantic aspects of their representations. In the first part of the sequence, they translate the phenomenon in inscriptions moving from the signified to the signifier. In the second part, when a new sign is introduced by the adults, they start moving from the signifier to the signified. In doing so, they engage in the two practices which together constitute a representational competence. The description of what has been observed and the rules by which the form of representation we have constructed operates evolve in unison and enrich our understanding of the "story of the balloon." The model is used to better make sense of the world. On the one hand, the Cartesian graph prototype the children constructed "forces" them into a linguistic formalisation: to establish a set of conventions that enables the whole group to construct a shared understanding of the signs used in it. On the other hand, as the ideas are gradually structured through words, in their concrete, abstract, metaphorical, analogical meanings, the children to realise their capacity to govern the relations between the signs and the world they represent. This is shown in the continuation of the discussion about the drawing.

A.: *It is as if it were a story, in the sense that this is a page that continues* [following the curve in the drawing his finger] *and here it is as if there were a full stop* [indicating the cross at 50], *and then it starts again here* [indicating the next cross] *and here the story ends* [indicating the cross at 80].

Teacher*: Why do you say that there is a full stop there?*

A.*: Because it's there that the balloon deflates.*

C.*: [looking carefully at the two graphs and running her finger along the curve asks A.] So this would be a whole page?*

A.*: No, it would be the whole story.*

Claudia nods, showing that she agrees with A.'s explanation.

F. *But then each cross would be a page?*

A.*: Yes!*

C.*: It's as if that page were the life of the balloon*.

In all classes, taking a cue from this work on the balloon, we then tried to introduce the use of a Cartesian graph as a tool to represent data collected the previous activities in which the children measured the weights and volumes of different materials. Significantly, in this activity the children used some of the key elements emerged in the sequence presented above to make sense of the graph in the new phenomenological context. In this case, time is not one of the variables included in the model. Nevertheless, the children pointed out that also this representation fulfils the function of telling a "story" of what has been observed. In particular, in commenting on a weight vs. volume graph obtained by measuring various quantities of salt in which one of the "experimental points" is decidedly not aligned with the others, A. points out that "to be a real story, the line [the one that was drawn by joining the points] should be straight." As one of his classmates points out that "in the balloon stories, there lines were curved," A. replies that "this story doesn't end, because I can weigh all the quantities I want." By saying so, he seems to refer to the fact that the correlation between the variables weight and volume established by the representation is a time-independent correlation. By drawing lines (and thus trying to tell stories, to take configurations, to construct models) that connect points (our observations of the reality around us, our experimental data), the children have come to discuss in a form that is certainly still basic the relationship between observations and model: the model (as well as a story) is "satisfactory" if it respects our idea of verisimilitude (to be a "true" story, as the children said) while a the lack of verisimilitude calls for a new negotiation of the meanings imbued in the model itself. A kind of reasoning that closely resembles Bruner's concept of "narrative thinking" as a complement and in many ways a precursor of "scientific thinking" (Bruner, 1996).

Conclusions

We have presented the findings of a long-term empirical study that involved ca. 100 children in the first grades of primary school as they engaged in modeling activities by making sense of phenomena and their relationship with different forms of representations. In order to do justice to the richness and complexity of the learning processes and the educational designs under consideration, we presented two extensive sequences that we consider exemplary of our findings. The children were exposed to an inquiry-based pedagogical design in which the adults supported the learning processes through scaffolding while avoiding direct instruction on models, representations and their relationship with phenomena. As research literature points out (see, e.g., Lehrer & Schauble,

2000), children should be given opportunities to develop their representational artifacts themselves as "the need for them has 'ripened' in the classroom and is widely apparent" (p. 46). Providing the children representation forms directly my result, as Lehrer and Schauble suggest, in manipulation of the symbols without understanding. We used, for example, whole class discussions as a fundamental moment of verbalization and sensemaking which can scaffold the development of understanding of representations and their relation with the represented phenomena. In semiotic terms this refers to the relation between signifier und signified which is central to the function of representations. Verbalization helps children become aware of this relation and fosters therefore also their metamodeling competence. The children involved in our activities spontaneously introduced a variety of symbols, metaphors, analogies and arguments to make sense of the syntactic and semantic aspects of their representations and negotiate the related epistemic norms. The sequences presented in this study exemplify a number of learning processes observed during our research programme that suggest that children in the first grades of primary school are able, when the educational activities are appropriately designed, to learn along all three dimensions of modeling competence: modeling practice, modeling product and metamodeling knowledge. This calls for reconsidering some assumptions on students' performance expectations as proposed in the *Framework for K-12 science education* (NRC, 2012) and the *Next Generation Science Standards* (NGSS Lead States, 2013).

References

Bruner, J. (1996). *The culture of education*. Harvard University Press.

Chiu, M.-H., & Lin, J.-W. (2019). Modelling competence in science education. Disciplinary and *Interdisciplinary Science Education Research, 1*, 1-11.

Cobb, P., Confrey, J., diSessa, A.A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9-13.

diSessa, A.A., & Cobb, P. (2004). Ontological innovation and the role of theory in design experiments. *Journal of the Learning Sciences, 13*(1), 77-103.

Hsieh, H.F., & Shannon, S.E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*(9), 1277-1288.

Kenyon, L., Schwarz, C., & Hug, B. (2008). The benefits of scientific modeling. *Science and Children, 46*(2), 40-44.

Lehrer, R., & Schauble, L. (2000). Developing model-based reasoning in mathematics and science. *Journal of applied developmental psychology, 21*(1), 39-48.

Lehrer, R., & Schauble, L. (2015). The development of scientific thinking. In L. S. Liben, U. Müller, & R.M. Lerner (Eds.), *Handbook of child psychology and developmental science: Cognitive processes* (pp. 671-714). Wiley.

Lichtman, M. (2013). *Qualitative research in education: A user's guide* (3rd ed.). SAGE Publications.

Mayring, P. (2020). *Qualitative Inhaltsanalyse*. Springer Fachmedien.

National Research Council (NRC). (2012). *A Framework for K-12 Science Education*. National Academies Press.

NGSS Lead States (2013). *Next generation science standards: For states, by states*. National Academies Press.

Nicolaou, C.T., & Constantinou, C.P. (2014). Assessment of the modeling competence: A systematic review and synthesis of empirical research. *Educational Research Review, 13*, 52-73.

Nielsen, S.S., & Nielsen, J.A. (2021). A competence-oriented approach to models and modelling in lower secondary science education: Practices and rationales among Danish teachers. *Research in Science Education, 51*, 565-593.

Odden T.O.B. and Russ R.S. (2019). Defining sensemaking: Bringing clarity to a fragmented theoretical construct. *Science Education, 103*(1), 187-205.

Passmore, C., Stewart, J., & Cartier, J. (2009). Model-based inquiry and school science: Creating connections. *School Science and Mathematics, 109*(7), 394-402.

Schwarz, C.V., Reiser, B., Davis, E., Kenyon, L., Acher, A., Fortus, D., Shwartz, Y., Hug, B., & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching, 46*(6), 632-654.

Siraj-Blatchford, I., Sylva, K., Muttock, S., Gilden, R., & Bell, D. (2002). *Researching Effective Pedagogy in the Early Years (REPEY) DfES Research Report 365*. Queen's Printer.

Upmeier zu Belzen, A., van Driel, J.H., & Krüger, D. (2019). Introducing a framework for modeling competence. In A. Upmeier zu Belzen, D. Krüger, & J.H. van Driel (Eds.), *Towards a competence-based view on models and modeling in science education* (pp. 3-19). Springer International Publishing.

A Teaching Innovation Proposal to Reduce Plant Blindness in Early Childhood Education

Beatriz García Fernández[1,2], Blanca Cherino Huertas[1], Esther Paños[3,2], José-Reyes Ruiz-Gallardo[3,2]

[1]University of Castilla-La Mancha, Faculty of Education of Ciudad Real, Castilla-La Mancha, Spain. [2]Botanic Institute. University of Castilla-La Mancha, Albacete, Castilla-La Mancha, Spain. [3]University of Castilla-La Mancha, Faculty of Education of Albacete, Castilla-La Mancha, Spain

Plants constitute an important content from the early years of school. This kingdom of living beings, as a part of the visible nature, is taught in pre-primary education and throughout compulsory education with different approaches and levels of complexity. However, despite their importance and presence in the education curricula, it is common that early childhood students do not recognize plants as living beings, which is known in the literature as plant blindness. This work presents a teaching innovation proposal designed to reduce plant blindness in early childhood education. It was implemented in a classroom of 3-year-old children where the topic of plants was traditionally approached using the textbook, so it was an innovative educational action. The effect of the proposal was assessed by applying an instrument before and after the intervention. Results show that the didactic sequence allowed to deepen the knowledge of plants in the short term.

Keywords: Plant blindness, early childhood education, science education

introduction

The content of living beings represents an important part of the educational curriculum. The current early childhood education curriculum addresses the subject of living beings within a framework of knowledge and approach to the natural environment, with special mention of respect and care for nature (Royal Decree 95/2022). In this sense, living beings that are visible to the human eye are of special importance at this stage, and plants are part of this group, along with animals.

Despite their relevance, plants are not always recognized or given the importance they deserve. This phenomenon is known in the literature as plant blindness. It is defined as the inability to perceive plants in the environment and to recognize their importance and characteristics, a phenomenon associated with the perception of

plants as inferior to animals (Wandersee & Schussler, 1999).

Some authors have pointed out some reasons for the existence of plant blindness. Strgar (2010) and Allen (2003) suggested that the human brain easily perceives elements that move and produce sounds around us, and that animals have qualities such as movement, feeding and ways of communicating that make them more visible than plants. Allen (2003) also points to the fact that plants do not constitute a risk to humans and, therefore, the human brain does not pay the same attention to plants as it does to animals.

Other authors, such as Balding and Williams (2016) reflect that the way formal education approaches plants also influences this phenomenon. In this vein, the teaching of plants may be seen by teachers as less interesting content than the teaching of animals (Prokop et al., 2010), which may lead to perceive plants solely as food and home for animals (Bermudez & García, 2015).

Previous works have identified plant blindness from university (Batke et al., 2020) to pre-primary education (Comeau, 2019; Paños et al., 2021) and have claimed for the need of teaching approaches to reduce this fact at school (Jose et al., 2019). In this context, the current work aimed at designing and assessing an innovation teaching approach for pre-primary education in order to reduce plant blindness in early childhood.

Method

It is a work framed in a mixed paradigm, with quantitative and qualitative data.

Sample

The sample was composed by 24 children (9 girls and 15 boys), with a mean age of 3,16 years old, and a SD of 0,37 years old.

The teaching innovation proposal sequence

The teaching innovation proposal designed included the following steps:

1. Presentation of eleven plant species using flashcards. Plants with flowers and other elements that attracted the children's attention, such as smell and morphological characteristics, were selected. These plants were: aloe vera (*Aloe barbadensis*), cactus (*Cynara cardunculus*), lavender (*Lavandulae angustifolia*), mint (*Menta spicata*), ranunculus (*Ranunculus asiaticus*), daisy (*Bellis perennis*), rose (*Rosa sp*), dahlia (*Dhalia officinalis*), gazania (*Gazania splendens*), primrose (*Primula vulgaris*) and carnation (*Dianthus caryophyllus*).

Figure 1. Plants used in the teaching intervention: a) Gazania, rose, mint, ranunculus, dahlia, daisy and lavender. b) Primrose and carnation.



2. Presentation in the classroom of the eleven real plants mentioned above (Figure 1).
3. Experimentation with plants: the children felt their scent and perceived their shapes and colors. They discussed their experiences with other classmates.
4. Participation in a card game in which they had to match pictures of the same plant species, mentioning its name. In this activity, the worked especially on the observation of the plants, since they had to look for similar key aspects in the images on the cards.
5. Identification of the names of the different plants worked on by means of a quiz.
6. Individual creation of a flowerpot with a sponge and printed images of the species worked previously (Figure 2).

Figure 2. Flowerpots stands made in the classroom using the species worked before.



Data gathering

Data was collected through an instrument composed by one graphic task ("Please, draw a plant") and two open questions ("Please, cite all the plant species you know", and "Please, cite all the animal species you know").

Data analysis

SPSS v. 22 software was used for data analysis. A descriptive analysis has been carried out, together to an inferential analysis. Due to the characteristics of the data, non-parametric statistics have been used: Wilcoxon test for related samples.

Results

Table 1 shows a summary of the results obtained in the two open-ended questions. The Wilcoxon test shows that the number of animals cited is slightly higher in the post-test than in the pre-test, but significant (Z = -6.041; p = 0.000). The number of plant species cited in the post-test is also significantly higher than in the pre-test (Z = -3.528; p = 0.000), but in this case the difference is more remarkable, rising the mean value from 0.29 to 4.79 (Table 1).

The results of the graphic task show greater detail in the children' drawings after the intervention, with significant differences in all the dimensions studied (Table 2). The number of dimensions analyzed in the drawings was limited due to the young age of participants.

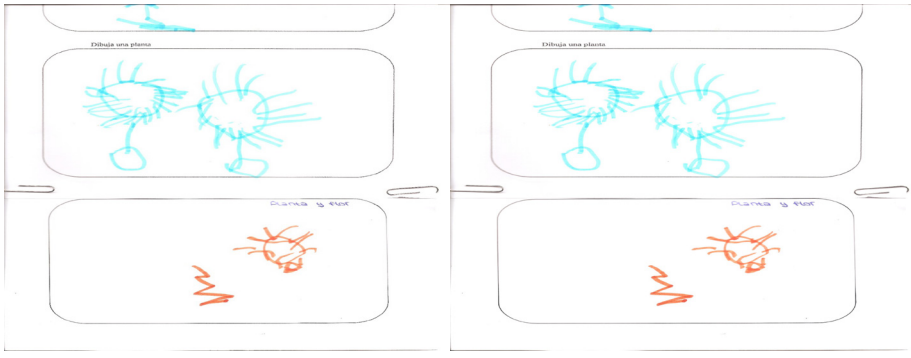
Table 1. Results obtained in the open question.

	Pre-test	Post-test	Z and p-value (Wilcoxon test).
Number of plant species cited	Media = 0.29 SD = 0.13	Media = 4.79 SD = 0.30	Z = -3.079 p = 0.002
Number of animal species cited	Media = 9.21 SD = 0.39	Media = 9.88 SD = 0.39	Z = -6.049 p = 0.000

Table 2. Results obtained in the graphic task.

	Pre-test (N=24)	Post-test (N=24)	Z and p-value (Wilcoxon test)
Flowers	12	23	Z = -4.796; p = 0.000
Tail	10	14	Z = -3.742; p = 0.000
Leaves	2	5	Z = -2.236; p = 0.000
Flowerpot	3	6	Z = -2.449; p = 0.000

Figure 3. Graphic production of the same child before (a) and after (b) the intervention.



Discussion and Conclusions

Plant blindness is a phenomenon well described in the literature, which points to the need to focus especially on plants from the early years of schooling. Thus, in response to this need, an innovative didactic approach has been designed and evaluated in a 3-year-old classroom. Due to the young age of the participants, data analysis focused on the number of plant species they knew compared to the number of animal species, and on the detail with which they were able to represent a plant. In both cases, the didactic intervention produced statistically significant improvements.

Children named more animal species than plant species. This is not surprising, considering that the human brain detects animals better than plants because of their movements and sounds (Strgar, 2010; Allen, 2003). Authors such as Prokop et al. (2010) point to more animal-centered didactic approaches as a cause of this phenomenon, but in the present study the didactic material has not been analyzed, and it is proposed as a future line of research.

After the intervention, the children were able to name more plant species than at the beginning. But what is most significant is the level of detail with which they elaborated their drawings, with more details of plant parts, especially flowers. This is not surprising considering that the flowers have been an important part of the didactic intervention. Since plants do not produce sounds or move like animals, the children's attention was attracted using other of their interesting features, such as flowers. Thus, it is not strange that they drew them more frequently than other plant elements in the post-test.The results show that the didactic proposal implemented has improved the knowledge of 3-year-old children about plants. However, this study has some limitations. Firstly, the data collection after the intervention was performed in the short term, so the impact should also be measured in the medium and long term. Secondly, it would also be interesting to design and evaluate a complementary intervention to help maintain the learning acquired with the proposal presented in this work. Finally,

this is a case study, and the results are not generalizable. Further research should be conducted to shed light on this topic.

Acknowledgements

The authors are grateful to the FEDER funding received by the University of Castilla-La Mancha in the frame of the project with reference 2022-GRIN-34471.

References

Allen, W. (2003). Plant Blindness. *BioScience, 10*(53), 926. https://doi.org/10.1641/0006-3568(2003)053[0926:PB]2.0.CO;2

Balding, M. & Williams, K.J.H. (2016). Plant blindness and the implications for plant conservation. *Conservation Biology, 6*(30), 1192-1199. https://doi.org/10.1111/cobi.12738

Batke, S.P, Dallimore, T. & Bostock, J. (2020). Understanting Plant Blindness – Students' Inherent Interest of Plants in Higher Education. *Journal of Plant Science, 4*(8), 98-105. https://doi.org/10.11648/j.jps.20200804.14

Bermúdez, J.M.M., & García Capocasa, C. (2015). La enseñanza de las plantas como un obstáculo educativo y los caracteres de visibilidad ecológica que pueden ayudar a superarlo: ¿qué especies consideran nativas los estudiantes de Córdoba? En G.M.A. Bermudez, y De Longhi, A.L. (Coordinadores), *Retos para la enseñanza de la biodiversidad hoy. Aportes para la formación docente* (pp. 293-325). Córdoba: Universidad Nacional de Córdoba. 2015. ISBN: 978-987-707-003-3

Comeau, P., Hargiss, C.L.M, Norland, J.E, Wallace, A. & Bormann, A. (2019). Analysis of Children's Drawings to Gain Insight into Plant Blindness. *Natural Science Education, 1*(48), 1-10. https://doi.org/10.4195/nse2019.05.0009

Jose, S.B., Wu, C.H., & Kamoun, S. (2019). Overcoming plant blindness in science, education and society. *New Phytologist Foundation, 3*(1), 169-172. https://doi.org/10.1002/ppp3.51

Paños, E., Pérez-Arroyo, E. & Reyes Ruiz-Gallardo, J. (2021). Las plantas como recurso educativo en el aula de Infantil. Intervención con plantas aromáticas. *Boletín de la AIA-CTS, 5*, 2183-5098.

Prokop, P., Prokop, M. & Tunnicliffe, S.D. (2010). Is biology boring? Student attitudes toward biology. *Journal of Biological Education, 1*(42), 36-39. https://doi.org/10.1080/00219266.2007.9656105

Royal Decree, 95/2022. Real Decreto por el que se establece la ordenación y las enseñanzas mínimas de la Educación Infantil. *BOE*-A-2022-1654.

Strgar, J. (2010). Increasing the interest of students in plants. *Journal of Biological Education, 1*(42), 19-23. https://doi.org/10.1080/00219266.2007.9656102

Wandersee, J.H & Schussler, E.E. (1999). Preventing Plant Blindness. T*he American Biology Teacher, 2*(61), 82-86. https://doi.org/10.2307/4450624

Early STEM Education: A Catalyst for Science Literacy in Preschool Children

Moleboheng Mokhele-Ramulumo

Department of Science and Technology Education, College of Education, University of South Africa, South Africa

This study investigates the impact of early STEM (Science, Technology, Engineering, and Mathematics) education on the science literacy of preschool children. It aims to assess disparities in science literacy between children exposed to STEM education and those following traditional approaches during the crucial early years of learning. Understanding how early STEM education influences the evolving educational landscape is crucial. Identifying gaps in science literacy between STEM and non-STEM preschoolers is crucial for informed educational practices. Recognizing the growing importance of STEM skills in the Fourth Industrial Revolution, early exposure to STEM education becomes essential. Aligned with the post-positivist paradigm, the study emphasizes the integration of diverse research methods and epistemological pluralism. Employing an embedded mixed-method research design, the study combines quantitative science literacy tests with qualitative semi-structured interviews. Through data triangulation, which combines quantitative science literacy tests with qualitative semi-structured interviews, the study comprehensively understands the research questions. Quantitative analysis reveals consistent STEM superiority in life sciences, earth sciences, and physical sciences. Qualitative insights uncover nuanced aspects of children's understanding, highlighting the depth of comprehension in both STEM and non-STEM groups. The study discusses the implications of the findings, emphasizing the significance of early STEM education in shaping a comprehensive understanding of science concepts. Disparities in science literacy underscore the need for tailored educational approaches to bridge gaps between STEM and non-STEM groups. The findings underscore the necessity of prioritizing STEM education in early childhood curricula to foster science literacy. Policymakers and educators can leverage these insights to enhance curriculum design, allocate resources, and provide professional development opportunities, ensuring a holistic integration of STEM education.

Keywords: STEM education, science literacy, early childhood

Introduction

In response to the rapidly evolving world, characterized by technological advancements and global challenges, there is a heightened global awareness of the imperative to equip the next generation with essential tools for success. Early

childhood science education, as highlighted by Wolff (2020) and Worth (2010), has emerged as a powerful catalyst in achieving this goal. It not only fosters intellectual growth and sparks curiosity but also plays a pivotal role in preparing children for future challenges. The nurturing of curiosity through exploration serves as the foundational step toward instilling a lifelong passion for science. This process also encourages future STEM learning.

Building on the understanding highlighted by Wolff (2020) and Worth (2010), the significance of science literacy has gained prominence. As argued by Coll and Taylor (2009), science literacy extends beyond mere factual knowledge, encompassing the critical understanding and evaluation of scientific concepts. Kaya et al. (2017) emphasize the essential nature of prioritizing science literacy from a young age, enabling students to fully comprehend scientific information and navigate the complexities of the modern world.

In countries such as Botswana in Africa, early childhood education is recognized for its pivotal role in fostering science literacy among children, as highlighted by Bose et al. (2013). Conversely, concerns arise in South Africa where a delayed introduction to science education until the fourth grade, as noted by James (2012), poses challenges affecting academic performance and shaping negative attitudes toward science (Reddy et al., 2013; Wellington & Ireson, 2008; Lynos, 2006).

The rationale for this study is firmly rooted in the global acknowledgment of the vital role played by early childhood science education in cognitive development (Darling-Hammond et al., 2019). Early exposure to science education, as emphasized by Darling-Hammond et al. (2019), is crucial for developing the skills necessary to tackle the challenges of the 21st century. The primary objective is to nurture curiosity, ignite a passion for science, and establish the foundation for future scientific pursuits and careers (Falk & Dierking, 2019).

This study aims to underscore the significance of early science education in fostering science literacy and sustaining interest in science. Therefore, by addressing gaps in South African early childhood science education, the study advocates for prioritization, hypothesizing that early exposure leads to higher science literacy levels. The central research question is: *To what extent does introducing science to children in early childhood impact their science literacy levels?*

Theoretical Framework

This study is firmly grounded in Gelma's (1990) Post-Piagetian Theory of Cognitive Development, highlighting children's robust cognitive capabilities crucial for science education (Kuhn & Pearsall, 2000). According to the theory, children possess domain-specific knowledge that allows them to develop

specialized understanding, including in science, through activities like play (Tricot & Sweller, 2010).

In building on this understanding, the significance of science literacy has gained prominence. This knowledge forms the basis for engaging in scientific inquiry skills, enhancing their abilities to observe, predict, and interpret data (Hatano and Inagaki, 1994).

In contrast to Jean Piaget's perspective, which asserts that children lack the abstract thinking necessary for science learning (De Luca, 1974), the Post-Piagetian viewpoint contends that children can comprehend fundamental scientific concepts within their cognitive abilities. Aligned to foster science literacy, this theory emphasizes the significance of early science education tailored to children's cognitive capacities and existing knowledge structures (Mnguni, 2014). It underscores children's readiness to learn science by nurturing their natural curiosity through age-appropriate educational interventions (Lindholm, 2018).

In conclusion, this theoretical framework serves as a guiding principle for exploring the relationship between early exposure to science education and science literacy levels among children. Therefore, by acknowledging and leveraging children's domain-specific knowledge structures, early science education is posited to facilitate the development of scientific understanding, laying a foundation for future advancements.

Methodology

The selection of post-positivism for this study is underpinned by the recognition of the importance of empirical evidence and the acknowledgment of subjectivity within research (Panhwar et al., 2017; Guba & Lincoln, 1989). Post-positivism, as a response to the objective reality stance of positivism, is particularly relevant to this research question as it allows for the exploration of subjective experiences alongside empirical evidence.

The choice of an embedded mixed-method research design is driven by the need for a comprehensive approach to address the research questions effectively. This design offers distinct advantages by allowing the integration of both quantitative and qualitative data, providing a more nuanced understanding of the relationship between early science education and science literacy. Qualitative data not only serves to confirm quantitative findings but also enriches the overall understanding through explanatory insights.

The study employed a two-phase data collection approach. During the initial phase, the quantitative research component involved the administration of a science literacy test for children aged 4 to 5, evaluating their proficiency in Life Sciences, Natural Sciences, and Earth Sciences. Rigorous development

procedures, including validation by experts and a pilot group, yielded favourable results, with an 82% content validity index and a reliable Cronbach alpha coefficient of 0.72. The meticulously crafted test was subsequently administered to 121 STEM school children and 87 non-STEM school children in Bloemfontein, South Africa, adhering to stringent matching criteria to ensure consistency and internal validity. To enhance reliability, children received guidance from teachers during the test, fostering a conducive and equitable testing environment. Ethical clearance was obtained to uphold responsible research conduct and prioritize child protection throughout the entire process. In the second phase, qualitative research was conducted through semi-structured interviews with six preschool children, drawn from both STEM and non-STEM schools, who had previously participated in the quantitative phase. These interviews were designed to delve into the thought processes and motivations underlying the children's responses to the test. To facilitate active engagement, the interviews incorporated the use of flashcards. The semi-structured format was chosen to provide flexibility for exploring individual perspectives and reasoning, allowing for a more comprehensive understanding of the qualitative data.

In the quantitative data analysis section, the specific statistical tests conducted in SPSS for comparing science literacy mean scores are detailed. This includes information on the analytical process, providing a clearer picture of the statistical methods employed. For qualitative data analysis, the deductive approach is elaborated upon, identifying predetermined categories and themes such as benefits of early science education, teaching challenges, and effective strategies. This approach was chosen for its suitability in aligning with the research objectives and applied systematically.

While ethical clearance is obtained, it is pertinent to outline the specific ethical considerations taken into account, particularly in the context of working with preschool children. This could include measures to ensure child protection, informed consent procedures, and the role of teachers in supporting children during the testing process. The methodology section concludes by summarizing the strengths and limitations of the chosen approach. Acknowledging potential biases or constraints ensures a transparent presentation of the research process. This reflective conclusion adds depth to the methodology, enhancing the overall credibility of the study.

Results

Results indicate that children exposed to science education demonstrate higher levels of science literacy. In Life Sciences, STEM children outperformed non-STEM peers by 4%, as evidenced by a significant t-test result ($t(105)=2.946$, $p<.001$). Similarly, in Earth Sciences, STEM children outperformed non-STEM peers by 8%, with a significant t-test result ($t(105) = 6.495$, $p<0.001$).

Additionally, in Physical Sciences, STEM children showed a slight advantage over non-STEM peers with a mean difference of 7%, supported by a significant t-test result ($t(105) = 3.917$, $p<0.001$).

Table 1. A comparison of science literacy levels between preschool children attending STEM and non-STEM schools

Construct assessed	Science domain assessed	*T*	*Df*	*P*	*MD*	Std. Error Difference	99% Confidence Interval of the Difference	
							Lower	Upper
Science literacy	Life	2.946	105	*p*<.001	4	1.3576	0.36	7.64
	Earth	6.495	105	*p*<.001	8	1.2315	4.66	11.34
	Physical	3.917	105	*p*<.001	7	1.7873	1.67	12.33

The results suggest that early exposure to STEM education has a positive impact on the development of science literacy in children. STEM children consistently outperformed their non-STEM peers in Life Sciences, Earth Sciences, and Physical Sciences, yielding statistically significant results. These outcomes underscore the importance of providing young children with opportunities for STEM education, positioning them for future success in academic and professional endeavours within STEM fields. It is noteworthy that the presentation focused on highlighting the most distinct results between STEM and non-STEM groups to enhance clarity regarding the differences in their ability to apply scientific concepts.

In the qualitative phase, regarding the Life Science item, children were prompted with the question: *Can you point out which things in the picture are alive (living) and which ones are not alive (non-living)?* Additionally, *can you explain what makes something alive or not alive? What special things do you look for to know if it's living or not*?

Figure 1. Life Science Task: Evaluating what children know about living and non-living things



The study's findings reveal a clear distinction in the comprehension of living and non-living objects between STEM and non-STEM children. When prompted to identify living and non-living things, STEM Child 1 and Non-STEM Child 1 provided insights that underscore these differences.

In response to the prompt, STEM Child 1 demonstrated a solid understanding of the difference between living and non-living things. They explained verbatim: *The dog, the boy, and the chicken are living because they can breathe and need water to grow*, highlighting the fundamental characteristics of living entities. This comprehensive grasp of basic biological concepts surpassed Non-STEM Child 1's awareness of a specific distinguishing feature. Non-STEM Child 1 suggested a more limited understanding of the broader characteristics that define living organisms, stating paraphrased: *The boy and the dog are living because the chicken cannot talk.* The findings indicate that STEM children possess a deeper understanding of the distinguishing features of living things.

For the Earth Sciences item, the children were asked, *can you think of things that people do that can harm nature? Can you explain why you think these things are bad for the environment?* Highlighting differences in environmental awareness, STEM Child 2 articulated a clear understanding of the detrimental effects of actions such as cutting down trees and factories making smoke. They expressed, *cutting down trees and factories making smoke is bad for us and the world. Trees give us fresh air; without them, it's hard to stay healthy. Factory smoke makes the air dirty, and it can make people, animals, and trees get sick*. STEM Child 2 emphasized the importance of trees in providing fresh air and the adverse impact of factory smoke on the environment, showcasing a deeper comprehension compared to Non-STEM Child 2's focus on the aesthetic aspects of environmental conservation by stating, *throwing papers on the floor makes our school not look nice*. The findings reveal significant disparities in environmental awareness between STEM and non-STEM children.

For the Physical Sciences item, the children were asked, *can you share your thoughts on whether the sun can make water in the outside tap warm? Please explain why you think so.* In examining the understanding of heat transfer, STEM Child 3 demonstrated a profound comprehension of the process by explaining, *yes, because the sun is hot. When the sun shines on the tap, it sends heat to the tap, and then the tap sends that heat to the water inside, making it feel warm.* On the other hand, Non-STEM Child 3 acknowledged the sun's ability to warm the water but provided a more straightforward explanation without delving into specific details about heat transfer, stating, *yes, because the sun is hot and will make the water in the tap warm.* These results underscore the notable differences in the understanding of scientific concepts related to heat transfer between STEM and non-STEM children.

In essence, the study illuminates significant disparities in the comprehension

of living and non-living objects, environmental awareness, and understanding of heat transfer between STEM and non-STEM children. STEM children consistently exhibit a deeper understanding of fundamental scientific concepts, emphasizing the potential benefits of integrating STEM education into early childhood curricula for a more detailed comprehension of scientific principles and natural phenomena.

Discussion

The study yields compelling evidence supporting the positive impact of STEM education on the science literacy of children. The results unequivocally demonstrate that children from STEM schools exhibit higher conceptual knowledge in science compared to their non-STEM counterparts. This underscores the pivotal role of early exposure to science curricula in fostering science literacy, aligning with Hyacinth's (2019) advocacy for prioritizing early STEM-based instruction over traditional approaches. The potential for improved science content knowledge through early science education is evident and resonates with Kusumastuti et al.'s (2019) research, confirming that integrated STEM curricula contribute to elevated science literacy levels in children. This convergence of results from various studies underscores the effectiveness of STEM education in promoting science literacy among young learners.

Building on these findings, the study places particular emphasis on the importance of Life Sciences education in developing children's comprehensive understanding of nature and living organisms. The findings support Demiral's assertion that children in STEM schools possess a more advanced understanding of life. This disparity is attributed to the challenges preschoolers face in grasping scientific concepts without meaningful connections to their environment, as noted by Harlen (2013).

Moreover, the study highlights the necessity of integrating environmental education into early childhood curricula. Aligned with Malebye's (2005) proposition that children exhibit a natural curiosity about the environment, the findings emphasize the potential of early environmental education to instil a sense of environmental responsibility. As Otto et al. (2019) stress, attitudes and behaviours acquired during early childhood significantly shape future environmental perspectives. The study's findings suggest that early environmental education can contribute to reducing pollution by enhancing environmental consciousness, echoing Duran's (2021) assertion that introducing environmental education at a young age can nurture heightened awareness of environmental concerns.

In addition to these insights, the study underscores the significance of providing opportunities for children to explore scientific concepts, fostering critical thinking skills and curiosity. Specifically, the investigation into children's understanding

of temperature changes resulting from heat transfer, with a focus on heat sources like the sun, revealed that STEM school children exhibited a more comprehensive understanding compared to their non-STEM counterparts. This aligns with prior research by Pahl et al. (2022) and Luce and Callanan (2020), indicating that preschool children often struggle to explain heat transfer concepts due to limited exposure. The study concludes that early science education plays a vital role in facilitating children's comprehension of abstract and complex phenomena during their formative years.

In conclusion, this study's findings contribute significantly to the discourse on early STEM education, emphasizing its positive influence on science literacy, the importance of meaningful connections in learning, and the potential benefits of integrating environmental education into early childhood curricula. These insights have implications for educators, policymakers, and researchers working towards enhancing science education in early childhood. Future research directions may explore effective pedagogical approaches for integrating STEM and environmental education, ensuring a well-rounded foundation for young learners.

Conclusion

In conclusion, this study underscores the critical need to prioritize and integrate STEM education into early childhood curricula to nurture science literacy from the earliest stages of development. Early childhood educators and curriculum designers emerge as pivotal contributors to this objective, responsible for crafting and implementing hands-on, inquiry-based activities that actively engage children in exploring and investigating scientific phenomena. These activities serve to foster critical thinking skills while deepening the youngsters' comprehension of scientific concepts.

To effectively implement STEM education in early childhood, policymakers should provide necessary resources and support, including funding for science materials and equipment, as well as offering professional development opportunities for educators. By investing in these vital aspects, policymakers guarantee that early childhood educators are well-equipped with the knowledge and skills necessary to seamlessly integrate STEM education into their curricula. This comprehensive approach not only empowers educators but also establishes the foundation for a scientifically literate generation. It prepares children for a future where science plays a central and influential role across various aspects of life.

References

Apps, J. N. (2008). Abstract thinking. In S. J. Loue & M. Sajatovic (Eds.), Encyclopedia of Aging and Public Health. Boston, MA: Springer Science & Business Media,

LLC.

Barz, C. L., & Achimas-Cadariu, A. (2016). The importance of developing scientific literacy skills: A case study. *Procedia - Social and Behavioral Sciences, 217*, 378-385. https://doi.org/10.1016/j.sbspro.2016.02.045

Bose, K., Tsamaase, M., & Seetso, G. (2013). Teaching of science and mathematics in pre-schools of Botswana: The existing practices. *Creative Education, 4(7)*, 43-51. https://doi.org/10.4236/ce.2013.47A1006

Bowman, B., Donovan, M. S., & Burns, M. S. (2001). *Eager to learn: Educating our preschoolers.* National Academy Press.

Coll, R., & Taylor, N. (2009). Scientific Literacy and Purposes for Teaching Science. *International Journal of Environmental & Science Education, 4(3)*, 289-299.

Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2019). Implications for educational practice of the science of learning and development. *Applied Developmental Science, 24(2)*, 97-140. https://doi.org/10.1080/10888691.2018.1537791

De Luca, F. P. (1974). Piaget and Science Teaching. *Iowa Science Teachers Journal, 11(4)*.

Duran, M. (2021). Perception of preschool children about environmental pollution. *Journal of Education in Science, Environment and Health (JESEH), 7(3)*, 200-219. https://doi.org/10.21891/jeseh.733800

Falk, J. H., & Dierking, L. D. (2019). Reimagining public science education: The role of lifelong free-choice learning. *Disciplinary and Interdisciplinary Science Education Research, 1(10)*. https://doi.org/10.1186/s43561-019-00010-3

Gelman, R. (1990). First principles organize attention to and learning about relevant data: Number and the animate-inanimate distinction as examples. *Cognitive Science, 14(1)*, 79-106. https://doi.org/10.1016/0364-0213(90)90027-T

Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation.* Sage Publications, Inc.

Hatano, G., & Inagaki, K. (1994). Young children's naive theory of biology. *Cognition, 50(1-3)*, 171-188. https://doi.org/10.1016/0010-0277(94)90027-2

Hyacinth, E. (2019). The Effect of STEM and non-STEM Education on Student Mathematics Ability in Third Grade (Doctoral dissertation). Walden University.

James, A. A., Beni, S., & Stears, M. (2019). Teaching science in the foundation phase: Where are the gaps and how are they accounted for? *South African Journal of Childhood Education, 9(1)*. https://doi.org/10.4102/sajce.v9i1.759

Kaya, E., Erduran, S., Akgün, S., & Aksöz, B. (2017). Nature of science in teacher education: A holistic approach. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education, 11(2)*, 464-501. https://doi.org/10.17522/balikesirnef.373423

Kuhn, D., & Pearsall, S. (2000). Developmental origins of scientific thinking. *Journal of Cognition and Development, 1*, 113-129. https://doi.org/10.1207/

S15327647JCD0101N_11

Kusumastuti, Y., Istiani, A., & Purnomo, C. (2019). Chitosan-Based Polyion Multilayer Coating on NPK Fertilizer as Controlled Released Fertilizer. *Advances in Materials Science and Engineering, 1-8*. https://doi.org/10.1155/2019/2958021

Lelliott, A. (2014). Scientific Literacy and the South African School Curriculum. *African Journal of Research in Mathematics, Science and Technology Education, 18(3)*, 311-323. https://doi.org/10.1080/10288457.2014.967935

Lindholm, M. (2018). Promoting Curiosity? Possibilities and Pitfalls in Science Education. *Science & Education, 27*, 987-1002. https://doi.org/10.1007/s11191-018-0015-7

Luce, M. R., & Callanan, M. A. (2020). Family Conversations About Heat and Temperature: Implications for Children's Learning. *Frontiers in Psychology, 11*, 1718. https://doi.org/10.3389/fpsyg.2020.01718

Malebye, V. E. (2005). Learners' understanding of the impact of air pollution on the environment in rural communities. Retrieved from http://hdl.handle.net/2263/27321

Mnguni, L. (2014). The theoretical cognitive process of visualization for science education. *Spring Plus, 3(184)*, 1-9. https://doi.org/10.1186/2193-1801-3-184

Otto, S., Evans, G., Moon, M. J., & Kaiser, F. G. (2019). The development of children's environmental attitude and behavior. *Global Environmental Change, 58*, 101947. https://doi.org/10.1016/j.gloenvcha.2019.101947

Pahl, A., Fuchs, H. U., & Corni, F. (2022). Young Children's Ideas about Heat Transfer Phenomena. *Education Sciences, 12(4)*, 263. https://doi.org/10.3390/educsci12040263

Reddy, V., Gastrow, M., Juan, A., & Roberts, B. (2013). Public attitudes to science in South Africa. *South African Journal of Science, 109(1-2)*. https://doi.org/10.1590/sajs.2013/1200

Ryan, A. B. (2006). Post-positivist approaches to research. In M. Antonesa, H. Fallon, A. B. Ryan, A. Ryan, T. Walsh, & L. Borys (Eds.), *Researching and writing your thesis: A guide for postgraduate learners* (pp. 12-28). Maynooth, Ireland: MACE, National Un.

Schulze, S. (2003). Views on the combination of quantitative and qualitative research approaches. *Progressio, 25(2)*, 8-20.

Spektor-Levy, O., Kesner Baruch, Y., & Mevarech, Z. (2013). Science and Scientific Curiosity in Pre-school-The teacher's point of view. *International Journal of Science Education, 35(13)*, 2226-2253. https://doi.org/10.1080/09500693.2011.631608

Tricot, A., & Sweller, J. (2013). Domain-Specific Knowledge and Why Teaching Generic Skills Does Not Work. *Educational Psychology Review, 26(2)*, 265-283. https://doi.org/10.1007/s10648-013-9243-1

Turiman, P., Omar, J., Daud, A. M., & Osman, K. (2012). Fostering the 21st Century Skills through Scientific Literacy and Science Process Skills. *Procedia* -

Social and Behavioral Sciences, 59, 110-116. https://doi.org/10.1016/j.sbspro.2012.09.253

Wellington, J., & Ireson, G. (2008). *Science Learning, Science Teaching.* London: Routledge. https://doi.org/10.4324/9780203821954

Wisdom, J., & Creswell, J. W. (2013). Mixed Methods: Integrating Quantitative and Qualitative Data Collection and Analysis While Studying Patient-Centered Medical Home Models. *PCMH Research Methods Series 13 (AHRQ Publication No. 13-0028-EF)*, 1-5.

Wolff, C. E. (2020). Early childhood science education: An important foundation for lifelong learning. *Early Child Development and Care, 190(10)*, 1586-1598. https://doi.org/10.1080/03004430.2019.1571067

Worth, K. (2010). *Science in Early Childhood Classrooms: Content and Process.* Cedar Falls, Iowa: SEED (STEM in Early Education and Development), University of Northern Iowa.



Strand 18

Teaching and Learning Science at Middle and Secondary School

Carla Morais & İbrahim Delen

Strand Chairs/Co-editors

Foreword

The teaching and learning of science at the middle and secondary school levels play a critical role in shaping the scientific literacy and curiosity of the next generation. It is within these formative years, covering grades 6 to 13, that students build a foundation of scientific knowledge, develop inquiry-based skills, and begin to appreciate the relevance of science in their daily lives and future careers.

Strand 18 of the ESERA 2023 proceedings, titled "Teaching and Learning Science at Middle and Secondary School (6-13 Grades)," delves into the diverse challenges, innovative strategies, and emerging trends that are defining science education in this crucial age group. The nine proceedings presented in this chapter offer a rich tapestry of research findings and practical insights that span a wide range of topics, from curriculum development and pedagogical approaches to the integration of technology and the fostering of scientific thinking among students.

These contributions highlight the importance of context and relevance in science education, showcasing how educators can inspire a deeper understanding and passion for science among students. Whether through hands-on experiments, cross-disciplinary projects, or the thoughtful incorporation of digital tools, the proceedings emphasize the need for dynamic and responsive teaching strategies that can adapt to the diverse needs of learners.

As you explore the research and discussions presented in this chapter, we hope you find valuable perspectives and ideas that can inform and enhance your own practice in science education. The insights gained here are not only a testament to the dedication and creativity of the researchers and educators who contributed but also a call to action for all involved in the education of future scientists, innovators, and informed citizens.

We are confident that the knowledge shared in this chapter will contribute significantly to the ongoing discourse on how best to engage middle and secondary school students in meaningful and effective science learning experiences.

Teacher Questions in Secondary Biology Classrooms: Making Scientific Practices Meaningful for Students

Zhongyan Zhang

School of Education, University of Leeds, UK

The purpose of this study was to explore how teachers used questions in secondary biology classrooms to engage students in scientific practices and how students perceived these questions. Six secondary biology teachers from four schools in Xi'an City in mainland China participated. Three lessons for each teacher were observed, followed by interviews in which the teachers were invited to reflect on their questioning practices. Eight students, aged 12 to 16, were selected from these classes and interviewed individually. The study found several key points. First, polished lessons that referred to the teacher's developing and repeatedly refining a lesson based on colleagues' suggestions after live classroom observation enabled more open questions. Second, teacher questions were driven by personal (e.g. teacher knowledge), internal (to the school), and external (e.g. national curriculum policy) factors, with external policy shift significantly affecting questioning practices. These factors either aligned to support the implementation of scientific practices or produced tensions that the teachers needed to work to negotiate. Third, the students demonstrated sophisticated and thoughtful reflections on, for instance, open and closed questions and scenario-based questions that asked them to think from a scientist's perspective. This expands our knowledge of students' views about teacher questioning. The findings have important implications for policy and practice. For example, teacher educators need to recognize the complexity of teacher questioning and its multiple factors to help teachers negotiate these tensions.

Keywords: Scientific practices; student voice; teacher questioning

Introduction

Scientific practices have been paid a great deal of attention by researchers, policymakers and practitioners in the realm of science education (e.g. Erduran, 2015; Osborne, 2014). Achieving curriculum standards and enacting these practices rely on "the teacher's ability to stimulate critical thinking skills through effective questioning behaviours" (Wilen, 1991, p.7). This study attempts to unpack how teachers use questions to engage students in the practices that 'scientists employ as they investigate and build models and theories about the world' (National Research Council, 2012, p.30) in secondary biology classrooms

in mainland China.

Teachers use questions to address their different purposes. Regarding the purposes of teacher questions, earlier research often relies on classroom observations to interpret them (e.g. Chen Hand, & Norton-Meier, 2017; Kawalkar & Vijapurkar, 2013). Thus, the teachers' accounts of their complex and manifold purposes are missing. Previous studies have primarily focused on the types of teacher questions, question techniques, and strategies (e.g. Blosser, 1975; Roth, 1996; Soysal, 2022). Thus far, however, there has been little mention of any controversy regarding teacher questioning. The impact of teacher training might be constrained if it solely aims to raise awareness among teachers about pedagogical questioning skills, without taking into account the tensions they experience. Thus, this paper intends to examine a large social and cultural context in which teacher questions are situated. This corresponds to Ryder (2015)'s research that classified 27 factors influencing teachers' responses to externally driven curriculum reform into three groups: personal (e.g. a teacher's subject knowledge and pedagogical skills), internal (e.g. students' differing backgrounds and aspirations and science department working practices), and external (e.g. national science curriculum reform). Exploring teacher questioning from personal, internal, and external perspectives contributes to an understanding of the complex reasons underlying their questioning practices. Furthermore, student voice can benefit not only teachers' professional development and school improvements (Flutter, 2007; McIntyre, Pedder, & Rudduck, 2005), but also enhance students' motivation, foster a sense of active community participation, and improve student attainment in science (Flutter & Rudduck, 2004; Rudduck & McIntyre, 2007). However, current research on students' views about teacher questions is limited.

This study has two principal aims. The first is to explore teachers' and students' reflections on teacher questioning, with the potential of addressing the research gap. The second aim involves exploring and exemplifying how teachers use questions to engage students in scientific practices in secondary biology classrooms to inform the design of teachers' professional-development activities and resources. The aims for the study can be shaped by the following research questions (RQs). In framing these RQs, the terms *teacher questions* and *teacher questioning* are distinguished to clarify the primary focus of each RQ. For some RQs, the term *teacher questions* was used. However, in others, questioning referring to strategies that encompass not only teacher questions but also consider their context and progression over time was emphasized.

- RQ 1 What types of questions do teachers ask during teaching that encourages students' engagement in scientific practices? This question uses quantitative data to describe the extent to which teachers pose different types of questions.
- RQ 2 What are the purposes of teacher questions when teachers engage

students in scientific practices in secondary biology classrooms? It explores the purposes of their questions, particularly those related to scientific practices. Here the term *teacher questions* was used. However, this RQ starts to shift its focus towards questioning strategies and the context of teacher questions. This is because the purposes of teacher questions often emerge from their setting and from the surrounding classroom discourse.

- RQ 3 What are teachers' reflections on factors influencing teacher questioning strategies? Teacher questions reflect teacher responses to national curriculum reforms, and they are influenced by not only individual characteristics but also by institutional and national context. This question takes a holistic perspective to explore the factors in teacher questioning practices on personal (e.g. teacher knowledge), internal (to the school), and external (e.g. national curriculum policy) levels.
- RQ 4 How do students perceive their teachers' questioning in class? This involves students' views on open and closed questions and what they notice about other aspects of teacher questioning.

Method

To address these questions, two main data-collection methods were adopted: lesson data and interviews. The formal data collection was conducted from March to November 2021 and originally aimed to collect data in their naturalistic setting and focus in depth on cultural context (Cohen, Manion, & Morrison, 2002). However, due to the global pandemic, the implementation of school visitations and lesson observations was not realised. Purposive sampling was undertaken to gain access to teachers who had the potential to use a relatively large number of dialogic and interactive questioning sequences (Mortimer & Scott, 2003). Six expert teachers participated in the study. *Expert teachers* here are defined as those who have substantial classroom experience in secondary biology and are recognized as exemplary teachers (e.g. recognition as a provincial expert teacher). The participant teachers exhibited a wide range of teaching experience, ranging from 10 to 26 years, across various grade levels and types of schools.

A sequence of three lessons and three interviews were conducted for each teacher. Each semi-structured, online interview lasted approximately one hour. Teaching episodes were used as an interview tool to help the teacher reflect on the lessons and elicit more ideas on scientific practices and teacher questions. A teaching episode here refers to an approximate 5-minute classroom teaching and learning sequence that mainly represents language exchanges between the teacher and students. The lessons are centered on scientific practices and encompass a variety of topics. These include the role of saliva, teeth, and tongue in digesting starch; healthy diet and food safety; exploration of the roles of

animals in the biosphere; the nervous system; the bones, joints, and muscles, and how they function in animal movement; the process of meiosis; the location of genes on chromosomes; and DNA as the main genetic material. Across the eighteen lessons, six lessons were polished lessons that referred to the teacher's repeatedly developing and refining a lesson based on colleagues' suggestions after their live classroom observation. This point is highly relevant to current policy developments in China. The term *polished lessons* is rarely used in the international literature, but it is the language used by Chinese teachers. Its meaning is deeply rooted in Chinese culture, where teaching is viewed as a craft. Teachers often dedicate significant effort and resources, such as physical models, lab equipment, and peer suggestions, to maximize the quality of their lessons. Lesson observation and discussion are integral components of this process. Generally, in the lesson discussion, the teacher who conducted the lesson begins by discussing the learning objectives, methods to achieve them, and strategies for tailoring teaching to students or educational theories. Subsequently, teachers who observed the lesson provide feedback on its strengths and weaknesses. This process may be repeated several times to achieve proximal teaching. After refining a lesson within a small group, the teacher presents it as a public lesson to a larger audience. Biology textbooks, slides, the videos teachers used in the classroom, photographs of lesson plans, physical models, lesson handouts, and students' work were valued and collected because they could provide important information about the context surrounding the questioning and teaching.

Eight students, aged 12 to 16 years, volunteered for the study following the distribution of information sheets and consent materials by their teachers. Recruitment focused on students from classes where lessons had been audio-recorded. All participant students provided informed consent and were happy to be involved in this study. These students were interviewed individually to gather their perspectives on teacher questioning, with each online interview lasting approximately 40 minutes. To gain a more in-depth understanding of the students' views, student drawings and teaching episodes were employed during the interviews. All lessons and interviews were in Mandarin and were audio-recorded and transcribed for analysis.

To present a broad picture of how teachers used different types of teacher questions, teacher questions were classified according to a framework adapted from Blosser's (1975) Question Category System for Science: closed, open, rhetorical, and managing discussion (see Table 1). To address RQ 2 and comprehensively understand teachers' actions and discussions, both lesson data and teacher interviews were utilised to determine the purposes of their questions. First, each teacher question was coded during the lesson in terms of its intended purpose from the author's perspective. Second, the interview data were used to develop codes for the different purposes of teacher questions from the teacher's

perspective. Third, the two strands of coding were compared and sorted into more general categories. Six phases of thematic analysis were employed for teacher and student interviews to address RQ 3 and RQ 4, respectively. The participants' stimulated recall of what they were thinking during classroom teaching and their comments on teacher questioning were analysed. The interview data were therefore paired with the lesson episodes to develop familiarity with the teaching and understand participants' perceptions. NVivo 12 was used to collate extracts for each code. To repeatedly engage with the data and develop deeper insights, coding entailed a continuous comparative analysis, involving iteratively reviewing, selecting codes, merging similar ones and discarding others.

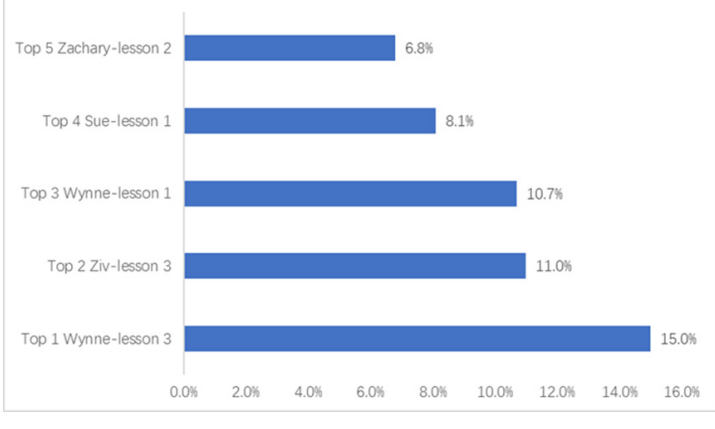
Table 1. Analytical framework for types of teacher questions, adapted from Blosser (1975)

Question type		Description	Example
Closed questions	Closed memory	Closed, to recall knowledge	1. Does it have ribosomes? 2. How many pairs of chromosomes does a somatic cell have?
	Closed other	Closed, to elicit higher cognitive thinking	1. How do you know it is starch? 2. Why was an equal mass of steamed buns added to each test tube?
Open questions	Open divergent	Open, a wide range of possible answers	1. Which activities can destroy the balance of an ecosystem? 2. Why cannot frogs survive in water-scarce terrestrial environments?
	Open evaluative	Open, to encourage students to elaborate from their own perspectives	1. Why do you think its skin has this function? 2. Have you ever felt…when did you use simple reflexes?
	Open scenario-based	Open, to think from another angle	1. Imagine you are the scientist Sutton or Morgan; what conclusions can you make from the data? 2. Imagine you are the spinal cord; what should you do if your finger is pricked, and neurons send this message to you?
Rhetorical questions		To emphasise an idea, to attract students' attention	1. Lack of water, isn't it? 2. S strain bacteria killed mice, right?
Managing discussion questions		To check consensus, to probe	1. Anything to add? 2. Does everyone agree with his idea?

Research Findings

Consistent with previous studies on teacher questions (Eliasson, Karlsson, & Sørensen, 2017), this study shows that open questions were significantly less frequent than closed ones. These two types of questions indicated teachers' power and control over broadening or narrowing student thinking. Teachers used too many closed-memory questions and students were not afforded sufficient opportunities to practice the use of scientific language and open up their thinking, which might have resulted in the authoritative classroom talk suggested by Mortimer and Scott (2013). Additionally, the context of polished lessons shows a higher proportion of open questions. Five lessons included a high proportion of open questions, and the top four lessons were polished ones (see Figure 1). Moreover, Wynne used more open questions and managing discussion questions than the other teachers and all three of her lessons were polished; this finding confirmed the association between open questions and polished lessons.

Figure 1. Top five lessons with a high proportion of open questions



(Note: Percentage=number of open questions in one lesson/total number of four types of questions in one lesson)

The purposes of teacher questions were inductively developed, including those related to scientific practices, the use of questions to emphasise social responsibility, use of questions to support the understanding of big ideas in science, and questions used for pedagogic purposes. For example, the pedagogic purpose of scenario-based questions, (e.g. "Imagine you are the scientist Sutton, what conclusions can you get from the data?") is to ask students to think in a particular situation. Some categories relate to the requirements of biology curriculum standards that emphasize the four dimensions of the core literacies: the big ideas of biology, scientific inquiry, scientific thinking, and social responsibility (Ministry of Education, 2017). Participant teachers purposefully implemented curriculum standards in their classrooms and used these ideas to explain their questioning and classroom behaviour.

Teachers' reflections on key factors in teacher questions were explored from three levels: personal, internal, and external. These findings show why a teacher asks a type of questions or a sequence of questions and what might be driving it, including, but not limited to, the external policy that requires conducting polished lessons, teachers' personal views about the openness of teacher questions, and their views about pre-planned and spontaneous questions. These factors align in some cases: teachers implement curriculum standards and motivate students to engage in scientific practices. For example, Wynne was encouraged to take part in a national competition; she selected a lesson about animal movements. She studied curriculum standards and identified the key teaching point of the lesson: that is, how muscles, bones and joints produce movement. Regarding the internal factor, one student expressed to Wynne confusion about how muscles, bones, and joints functioned even after the lesson. This feedback prompted Wynne to reflect on the shortcomings of the experimental material provided in the textbook and to reconsider her teaching approach. Recalling the classical practical work she had undertaken during her Master's degree studies, Wynne drew on this personal experience to revise the experimental material of her lesson to help students investigate how to generate movement. External, internal, and personal factors were interacting with each other towards the same direction: implementing curriculum standards and helping students draw a conclusion based on results.

However, in other cases, personal/internal/external factors push the teachers in opposite directions, which lead to teacher tensions. For example, teachers' desire to use more open questions conflicted with their task of supporting students in acquiring prescribed subject knowledge in large class settings. However, tension does not always lead to stress and being pulled in different directions can lead to generating positive and creative energy. For example, Sue devised a sequence of questions and classroom activities designed to efficiently utilize limited time for covering the teaching content, while also aiming to fulfil her personal goal of fostering scientific thinking and actively engaging students in scientific practices.

The students showed similar views in terms of different types of teacher questions (e.g. open, closed, and scenario-based questions) and they noticed the question content, pedagogic issue, and effect of teacher questioning. In the following example, a student stated that scenario-based questions pushed them to imagine that they were a scientist and to try to think and make decisions from the scientist's perspective: "I can clearly remember the moment when he taught about Mendel, he asked if you were Mendel, when he taught about Sutton, he asked if you were Sutton." Students also used a range of terms to describe their views on teacher questions, for instance, "challenging", "interactive", "developing thinking", "paving the way for later learning", "broadening knowledge", "using the occasion

to stop zoning out", "focusing on easily confused knowledge", and "increasing difficulty gradually". These examples show the students' sophisticated and thoughtful reflections on the pedagogic role of questions.

Discussion and Conclusions

The findings of this study confirmed the findings of previous research that teachers tend to use far more closed questions than open (e.g. Eliasson et al., 2017; Lee & Kinzie, 2012). Additionally, it was noted that teachers less frequently employed managing discussion questions that promote student participation in class discussions. However, this study also demonstrated that teachers used questions differently in two different class types and used more open questions in polished lessons.

The interview data and classroom observations were used to access the teachers' purposes when questioning, which involved teachers' clearly explaining their goals in the interviews and the author's postulations regarding the intentions of teachers' classroom behaviours. This approach was a distinctive one because previous studies tended to rely on classroom observations to approach and interpret the purposes of teachers' questions. The finding shows that teachers used a wide variety of purposes, involving cognitive, affective, and pedagogical domains, and that they considered inquiry skills, scientific thinking, scientific literacy, and the nature of science. This contradicts the stereotype of China being a country that focuses on knowledge recall, closed questions, and exam success. These focuses are present but stand alongside focuses on open exploration, value education, and the application of science ideas within society.

The use of teacher questions cannot be dissociated from other teacher practices; therefore, any consideration of influences on teacher questioning must consider factors more broadly (Boyd, 2015; Chin, 2007; Kayima & Jakobsen, 2020). Teacher questions depend on personal (e.g. teachers' views about teacher questions), internal (e.g. students' views and the school's working practices), and external (e.g. the high school or national college entrance exams) factors. The role of polished lessons is significant because most questions were planned and devised, and polished lessons effectively integrate personal, internal, and external factors, thereby encouraging teachers to refine both their questioning methods and overall teaching practices.

Students can present insightful reflections on the role of teachers' questioning. Their focus extends beyond the conceptual elements of these questions to encompass a range of domains, such as pedagogical, social, affective, and epistemic aspects.

Several recommendations regarding student noticing, policy makers, teacher developers, and teachers' professional development are discussed. For example,

teachers could clearly articulate what they did in the classroom and why they asked a particular question or sequence of questions. Such articulation represented a reconstruction of their thinking and a way of using language to develop thinking and then to develop teaching (Brown & McIntyre, 1993). An implication of this is the possibility that teacher educators use teaching episodes or videos to encourage teachers to talk about what they did and why they did it in an environment that emphasizes respect and expertise sharing (Zhu & Qin, 2008), rather than criticism.

Educators in teacher-training programs must recognise the complexity of teacher questioning that is driven by multiple factors and help teachers use questions to negotiate tensions in scientific practices. A case study of how an expert teacher has negotiated these challenges positively could be used. During teacher training sessions, the teacher educator could give teachers opportunities to talk about their challenges and pressures in their own context and how they overcame these challenges. Students pay attention to different aspects of teacher questioning, including the conceptual, pedagogical, social, and affective domains. Teacher educators can present views from students to teachers as a powerful tool for raising awareness. This approach helps teachers understand the various types of attention students exhibit and the messages those convey. It also encourages teachers to develop a diverse set of purposes for using questions. This method differs from those in previous studies that have focused on contrasting expert and novice teachers in professional-development programs (e.g. Carter, Cushing, Sabers, Stein, & Berliner, 1988; Jacobs, Lamb, & Philipp, 2010).

Acknowledgement

I would like to extend my sincere gratitude to Jim Ryder and Michael Inglis for their significant contributions to my PhD study. I am also grateful to Vanessa Kind for her encouragement and for listening to my presentation at the conference.

References

Blosser, P. E. (1975). *How to ask the right questions*. Washington, D.C.: National Science Teachers Association Press.

Boyd, M. P. (2015). Relations between teacher questioning and student talk in one elementary ELL classroom. *Journal of Literacy Research, 47*(3), 370–404.

Brown, S. A., & McIntyre, D. (1993). *Making sense of teaching*. Open University Press.

Carlsen, W. S. (1991). Questioning in classrooms: A sociolinguistic perspective. *Review of Educational Research, 61*(2), 157–178.

Carter, K., Cushing, K., Sabers, D., Stein, P., & Berliner, D. (1988). Expert-novice differences in perceiving and processing visual classroom information. *Journal*

of Teacher Education, 39(3), 25–31.

Chen, Y., Hand, B., & Norton-Meier, L. (2017). Teacher roles of questioning in early elementary science classrooms: A framework promoting student cognitive complexities in argumentation. *Research in Science Education, 47*(2), 373–405.

Chin, C. (2007). Teacher questioning in science classrooms: Approaches that stimulate productive thinking. *Journal of Research in Science Teaching, 44*(6), 815–843.

Cohen, L., Manion, L., & Morrison, K. (2002). *Research methods in education*. Routledge.

Eliasson, N., Karlsson, K. G., & Sørensen, H. (2017). The role of questions in the science classroom – how girls and boys respond to teachers' questions. *International Journal of Science Education, 39*(4), 433–452.

Erduran, S. (2015). Introduction to the focus on … scientific practices. *Science Education, 99*(6), 1023–1025.

Flutter, J. (2007). Teacher development and pupil voice. *The Curriculum Journal, 18*(3), 343–354. https://doi.org/10.1080/09585170701589983

Flutter, J., & Rudduck, J. (2004). *Consulting pupils: What's in it for schools?*. RoutledgeFalmer.

Jacobs, V. R., Lamb, L. L., & Philipp, R. A. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education, 41*(2), 169–202.

Kawalkar, A., & Vijapurkar, J. (2013). Scaffolding science talk: The role of teachers' questions in the inquiry classroom. *International Journal of Science Education, 35*(12), 2004–2027.

Kayima, F., & Jakobsen, A. (2020). Exploring the situational adequacy of teacher questions in science classrooms. *Research in Science Education, 50*(2), 437–467.

Lee, Y., & Kinzie, M. B. (2012). Teacher question and student response with regard to cognition and language use. *Instructional Science, 40*, 857–874.

McIntyre, D., Pedder, D., & Rudduck, J. (2005). Pupil voice: Comfortable and uncomfortable learnings for teachers. *Research Papers in Education, 20*(2), 149–168. https://doi.org/10.1080/02671520500077970

Ministry of Education of the People's Republic of China. (2017). *Biology curriculum standards for senior secondary education (2017 edition)*. People's Education Press.

Mortimer, E. F., & Scott, P. H. (2003*). Meaning making in secondary science classrooms*. Open University Press.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.

Osborne, J. (2014). Scientific practices and inquiry in the science classroom. In N. G. Lederman & S. K. Abell (Eds), *Handbook of research on science education*

Volume II (pp. 579–599). Routledge.

Roth, W. M. (1996). Teacher questioning in an open-inquiry learning environment: Interactions of context, content, and student responses. *Journal of Research in Science Teaching, 33*(7), 709–736.

Rudduck, J., & McIntyre, D. (2007). *Improving learning through consulting pupils*. Routledge.

Ryder, J. (2015). Being professional: Accountability and authority in teachers' responses to science curriculum reform. *Studies in Science Education, 51*(1), 87–120.

Soysal, Y. (2022). Science teachers' challenging questions for encouraging students to think and speak in novel ways. *Science & Education*, 1–41.

Wilen, W. W. (1991). *Questioning skills, for teachers. What research says to the teacher*. 3rd ed. National Education Association.

Zhu, X., & Qin, J. (2008). An investigation on the relationship between open class and teachers' professional development. *Curriculum, Teaching Material and Method, 28*(5), 83–88.

A Conceptual Framework for Digital Research Skills in Secondary Science Education

Kim Blankendaal, R. F. G. Meulenbroek, W. R. Van Joolingen

Utrecht University, Utrecht, Netherlands

This study focuses on the perceived gap between the required and actual level of digital research skills (DRS) of students entering tertiary science education. A framework of DRS, incorporating seven categories, is evaluated using an exploratory qualitative study employing semi-structured interviews with university teachers (N = 15). The level of DRS at the start of university science education and university teachers' perceptions of first-year students' level of DRS have been investigated. The results show that the skills of writing a research paper using digital tools, using proper resources, and analyzing, transforming, and visualizing data were generally found to be wanting.

Keywords: Secondary science education, digital research skills, transition gap.

Introduction

First-year university students have been reported to lack skills related to adapting, interpreting, and evaluating outcomes during data processing and working with statistics (Oakleaf & Owen, 2010). These results were mirrored in a study by Akuegwu and Uche (2019) reporting a low level of (general) RS in pre-university students: reading, presentation, communication and information gathering skills were found to be adequate, whereas data analysis was found wanting. Other research has confirmed this problematic level of skills such as information-seeking and using ICT among students in secondary education (Julien & Barker, 2009; Smith et al., 2013; Walraven et al., 2008). Despite the importance of DLS and RS for academic success (Warschauer et al., 2004), only a few scientific studies have addressed the development of these skills in secondary education and the way this affects the ensuing transition to science undergraduate education. A formal combination of RS and DLS was not been encountered in this search, and little appears to be known about the assessment of the combination of these skills in secondary education. The main aim of this study is (1) to provide a framework for the combination of DLS and RS, which we propose to label Digital Research Skills (DRS) and (2) to evaluate how proficient science university teachers judge their incoming students to be in these skills. Our research question is: Which digital research skills (DRS) show a gap between the final level of pre-university and the required entry level for

science studies, as perceived by university teachers? The three sub-questions are: (1) How can research skills and digital literacy skills be integrated in a usable combined framework for digital research skills (DRS)? (2) Which DRS do university teachers consider important for beginning science students? (3) What are university teachers' perceptions about first-year students' level of DRS?**4**

Research Design and Methods

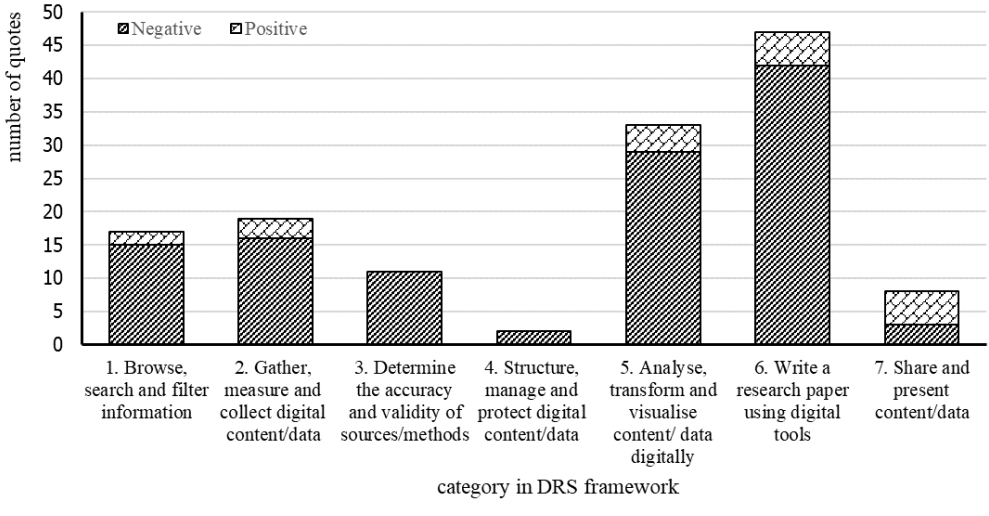
Due to the rapidly changing field, we have studied publications from the past ten years that were related to or assessed in secondary education. To answer the first sub-question, we have combined several leading frameworks for DLS with the RSD framework presented by Willison (2018) in order to arrive at a framework for DRS. We looked at research skills in the RSD framework that highly correspond to DLS in the other (ICT) frameworks. Our DRS framework therefore incorporates seven categories: 1. Browse, search and filter information; 2. Gather, measure and collect digital content/data; 3. Determine the accuracy and validity of sources/methods; 4. Structure, manage and protect content/data; 5. Analyze, transform and visualize content/data; 6. Write a research paper using digital tools and 7. Share and present content/data.

In order the answer the second and the third sub-questions, we conducted an exploratory qualitative study to clarify what DRS are required at the start of higher science education in our country. This qualitative study was based on semi-structured interviews with 15 academics who teach at the university level. The academics were professors, teachers, coordinators, program directors and education directors and were selected based on the following criteria: they teach first-year chemistry and/or physics students and supervise or teach these students when the students are during a research project. Out of a total of 30 candidates, 15 responded; the interviews were conducted on a voluntary basis. Participants provided written informed consent, which was repeated orally at the start of the interview. The academics were men and women in different age categories spread over eight theoretical and two technical universities in the country. All recordings were transcribed and pseudonymised. Potentially interesting quotes from the participants were selected and coded using the categories in the DRS framework. Similar skills were grouped together per category. Each quote was also coded as reflecting either a positive or negative perception by the interviewee. We used a inductive approach and looked at how often a skill was mentioned per category. Duplicate coding was done on 50 of the 223 quotes; Cohen's kappa was found to be 0.88 for determining whether a quote was positive or negative and 0.81 for assigning the codes to the seven categories of the DRS framework, indicating good to near-perfect interrater reliability.

Results

In Figure 1, the total number of quotes on the students' level of proficiency is displayed, divided into positive and negative quotes. A positive quote means that the interviewee was mostly satisfied with the extent to which students applied the corresponding skill. A negative quote means the interviewee was generally not satisfied. It is striking that most quotes (106 out of 125), in response to the third interview question were negative. The negative quotes mainly indicate a lack of skill in: writing a research paper using digital skills; analysing, transforming and visualising content/data; and gathering, measuring and collecting digital content/ data. The positive quotes mainly related to skills such as sharing and presenting content/data and also writing a research paper using digital tools.

Figure 1: Frequency of positive and negative interview quotes per category of our DRS framework (N = 15).



The majority of the negative quotes in category 6 'Write a research paper using digital tools' are related to big differences in skill level between the students (11 quotes given by 8 interviewees). Other quotes refer to lack of skill in using captions and references (10 quotes by 8 interviewees), line of reasoning in a paper (10/6) and text formatting (8/5). The positive quotes (5/3) related to general skills in writing a paper.

In addition to the lack of adding an explanation in the caption (14/8), a significant number of the negative quotes in category 5 'analyse, transform and visualise content/data digitally' are related to students' lack of digital skills to process data (8/5). Michael (Coordinator Chemistry Practical) said:

> 30 to 40 percent of the students can handle it just fine and others have never done it. So if you look in their lab journal on the first few experiments, they have a hand-drawn graph, while you really said, well, try doing it in Excel, but then they say yes, they have no idea how it should be.

Other quotes refer adding a correct axis label (8/8) and adding a trendline instead of connecting the measure points (7/7). The positive quotes were about visualizing content/data in general (4/3).

Conclusion and Discussion

By structuring and comparing both RS and DLS in pre-university education, we formulated seven categories in a usable framework for DRS. The 15 interviewed academic science teachers and professors generally agreed on the importance of writing skills using digital tools. The university teachers' perceptions were mostly negative about the students' DRS proficiency. The primary problem concerned the writing of a research paper using digital tools. Students lack the ability to structure a paper and lack skills in referencing, use of captions and formatting. Many students were found having insufficient experience with using spreadsheets, e.g., in processing data , axis labelling, drawing of a trendline, etc. However, the interviewees note big differences between individual students in this respect. Furthermore, students are reported to demonstrate difficulties browsing, searching and filtering information, consistent with Julian & Barker's (2009) and Walravens' (2008) observations. In contrast, the interviewees were mostly satisfied with students' level of sharing and presenting content/data.

Our research question was: Which digital research skills (DRS) show a gap between the final level of secondary education and the required entry level for science studies, as perceived by university teachers? Our results showed that there is indeed a gap in DRS between the final level of secondary education and the required entry level for science studies. The perceived gap is most pronounced in analysing, transforming and visualising content/data digitally and/or writing a research paper using digital tools.

The limitation of this study is the relatively small sample size and that it solely reflects the situation in the Netherlands. However, the general literature cited in the introduction showed that the issues addressed are encountered in other countries as well (Akuegwu & Uche, 2019). Students appear to be especially lacking in skill on data processing and writing a research paper. More attention to data processing and writing a research paper in secondary science education could help to bridge this gap. This could be reflected in the learning goals in the syllabi of science subjects in secondary education. Our study could provide a step towards identifying national and international digital research skill levels in order to bridge the gap from pre-university to academic science education. An important next step is to analyse to what extent skills from the DRS framework

are actually applied by pre-university students. assignments.

References

Akuegwu, B. A., & Uche, K. (2019). Assessing graduate students ' acquisition of research skills in universities in Cross River State Nigeria for development of the total person. Educational Review: International Journal, 6(1), 27–42.

Julien, H., & Barker, S. (2009). How high-school students find and evaluate scientific information: A basis for information literacy skills development. Library and Information Science Research, 31(1), 12–17. https://doi.org/10.1016/j.lisr.2008.10.008

Oakleaf, M., & Owen, P. L. (2010). Closing the 12 - 13 gap together: School and college librarians supporting 21st century learners. Teacher Librarian, 37(4), 52–58.

Smith, J. K., Given, L. M., Julien, H., Ouellette, D., & DeLong, K. (2013). Information literacy proficiency: Assessing the gap in high school students' readiness for undergraduate academic work. Library and Information Science Research, 35(2), 88–96. https://doi.org/10.1016/j.lisr.2012.12.001

Walraven, A., Brand-Gruwel, S., & Boshuizen, H. P. A. (2008). Information-problem solving: A review of problems students encounter and instructional solutions. Computers in Human Behavior, 24(3), 623–648. https://doi.org/10.1016/j.chb.2007.01.030

Warschauer, M., Grant, D., Del Real, G., & Rousseau, M. (2004). Promoting academic literacy with technology: Successful laptop programs in K-12 schools. System, 32(4), 525–537. https://doi.org/10.1016/j.system.2004.09.010

Willison, J. W. (2018). Research skill development spanning higher education: Critiques, curricula and connections. Journal of University Teaching and Learning Practice, 15(4).

Secondary Students Drawing Comics for Physics Learning

Evi Ketsea[1], Emmanuel Rollinde[1], Cécile de Hosson[2]

[1]CY Cergy Paris Université, LDAR, France

[3]Université Paris Cité, LDAR, France

The objective of this study is to explore the potential benefits of students drawing comics on the learning of physics in the context of the secondary classroom. A design-based research study comprises the design and teaching of an interdisciplinary (physics and arts) lesson sequence in a class of a French secondary school (ages 14-15). Analysis of the comics production will use semiotic and didactics theoretical frameworks.

Keywords: Physics, comics, school

Focus of the Study

The objective of this study is to explore the potential benefits of students drawing comics on the learning of physics in the context of the secondary classroom.

A widely accepted definition of comics is "juxtaposed pictorial and other (including letters and words) images in deliberate sequence" (McCloud, 1993, p.9). Moreover, comics are considered as using a visual language that has its own lexicon (comprising items of various morphologies), grammar, and including a narrative, a navigational and a conceptual structure (Cohn, 2013). As such, the language of comics lends itself to the hypothesis that it can be utilised in teaching and learning and especially in physics. The "cohabitation of words and pictures in a sequential-simultaneous ecosystem" (Sousanis, 2015, p. 64), appears to be particularly conducive for the cognitive manipulation of numerous physics concepts such as changes in the points of view (and reference systems or standards), changes in scale/scope of different interactions, the formalism/symbolism as for example force vectors, field lines and diagrammatic representations, to name some examples.

For example, in the case of causal reasoning and simultaneity (often related to students misconceptions in physics), previous research (Viennot, 2007) has shown that when students formulate explanations in a narrative form, they often adopt a common reasoning characterised by linear simplification, presenting a chain of transformations related to only one of the variables involved (with an underlying chronology which turns the explanation into "a story"). In this case, the comics format can deliver more sophisticated possibilities (e.g., representing

causality for one variable and simultaneity for others) because although comics are read sequentially like text, the entire composition is also viewed all at once as a connected space, a web of elements forming a cohesive whole, where there is a spatial interplay of the sequential and the simultaneous (Sousanis, 2015).

Review of Literature

Research on the (various ways of the) use of comics in the classroom started in the 90s (parallel to certain acclaimed publications on the subject of comics itself - e.g., McCloud, 1993 - that contributed to raising its previously lower status of a popular art form of poor value). In general, these research studies review and analyse published comics aiming at communicating science to the public (including school students), identify their advantages and constrains (Tatalovic, 2009) or focus on the science popularisation characteristics adopted (Baudry & Crepin, 2019; Farinella, 2018).

One of the ways to use comics in the classroom is the creation of comics by the students. Here, the scope of research is more limited and the part focusing on secondary science education even more so. Gonzales-Espada (2003) presents a teaching strategy of having students draw "a scientifically accurate comic strip" in a physics class. In Lo Iacono et al (2011), comics by students in a biology class aimed at raising student's awareness on Nature-of-Science concepts and the process of peer-review. In a study by Albrecht et al (2012), students drew comics after a period of physics teaching for learning outcomes evaluation. The one clear general result in all the above studies was that those activities increased student motivation. Finally, in a research project by de Hosson et al (2019), teenagers created comic pages on science and maths topics priorly presented by a scientist and the comics produced were analysed in terms of the relation between the science content and the choices of the comics' narrative and graphical characteristics/elements employed.

Although it has been widely claimed that drawing comics can enhance learning by promoting multi-modal thinking, ideas generation, dealing with ambiguity and making connections (Carlson et al, 2020; Sousanis, 2015), these questions have not been addressed in research. Also, more specifically, the links between the activity and the actual students' learning (outcomes or processes) rest unexplored.

Research Questions

- R.Q.1: What aspects of the students learning in physics are enhanced due to their drawing comics?
- R.Q.2: What are the characteristics of the comics creation activity that enhance each aspect of learning and what are the processes through which

this is achieved?

Theoretical Framework

We are planning to use three theoretical frameworks.

Firstly, the TAD theory (Chevallard, 2013) will be used for both the design and the analysis in this research on a macro to a meso level comprising all the components that make up the researched instance of physics education ("didactique de la physique"): institutions, actors, actions, tasks, skills, knowledge, and theories. Through the structure provided by this theory, we will look at:

> 1) The effects on learning due to the functioning of the system ("system didactique") consisting of all the above-mentioned components.
>
> 2) The "praxeologies" ("packets" of tasks, skills and the -science education- reasoning that justifies and relates them to a specific physics knowledge) employed and their effect on learning.

In a second, meso-micro level, this study will look closer at the learning processes at play during the comics creation by employing the theoretical framework of the "instrumental genesis" of Rabardel (1995). In this case, the comics can be viewed as a "psychological instrument" (as do language, symbols, and diagrams according to the Vygotskian tradition) and also as a "semiotic instrument" (Rabardel, 1995) with its own visual language. This theory views "an instrument" as including an artifact/object component (in our case: the comics pages) and a "utilisation scheme" component comprising the actions taken during its use (creation/development). In our case, these actions correspond(/reflect) to the learning processes.

In a third micro level, we will link the comics creation outcomes/processes to the learning/cognition ones using semiotic theory. This theory views both learning and cognition as a semiosis i.e., the processes through which significations emerge, transform, evolve, and disappear (Cunningham, 1998). Analysis of the signs (and their evolution during the creation process) in the comics will provide a mapping to the learning and cognition processes at play. Shank's classification of signs (Cunningham, 1998) adapting Peirce's classifications in an educational context will be a valuable tool here.

Research Design and Methods

The crossdisciplinarity in this project lends itself to a Design-Based Research (DBR) methodology as it is "a systematic but flexible methodology aimed to improve educational practices through iterative analysis, design, development, and implementation, based on collaboration among researchers and practitioners in real-world settings, and leading to contextually-sensitive design principles

and theories" (Wang & Hannafin, 2005, p.6).

The DBR experiment took place in collaboration with a physics and an arts teacher, in a class "3ème du college" of a French secondary school (ages 14-15) with 29 students. It included 6 physics and 4 arts lessons. The physics learning objectives were the ones concerning the whole unit "Forces and Interactions" as they were described in the Program issued by the Ministry of Education.

In this methodological context we first designed a lesson sequence comprising physics and arts lessons (on drawing/comics creation). In the second phase the implementation of the lesson sequence was accompanied by the data collection. Our primary data corpus is composed of the series of produced comics by every student. Additionally, there are audio recordings of the teachers' discourse.

We are currently performing the qualitative data processing by analysis including:

1) Identification of the "praxeologies" (Chevallard, 2013) packets made possible by the activity of comics creation.

2) Identification of the "utilisation schemes" (Rabardel, 1995) employed by the students mapping the learning processes along the shaping of the final state of the comics instrument.

3) Semiotic analysis of the finished and evolving comics identifying the dimensions /characteristics of the learning processes (types of reasoning, concepts representation and processing, problem-solving strategies, interactions between representation choices, metacognition, etc.) and their evolution/development, based on the signs and the semiotic characteristics found in the comics and the visual language employed.

References

Albrecht, E. and Voelzke, M. R. (2012). 'Creating comics in physics lessons: an educational practice'. *Journal of Science Education*, 2 (13), pp. 76–80.

Baudry, J. et Crépin, O. (2019). La Fabrique de recherche Dessinée. *Telling Science, Drawing Science* (Conference) May 2019, Angoulême, France.

Carlson P., Garcia A., and Kirtley S. (editors) (2020). *With great power comes great pedagogy*. University Press of Mississippi, 2020

Chevallard, Y. (2013). Éléments de théorie anthropologique du didactique (TAD) : Une initiation à la didactique fondamentale. *Journée du mercredi 9 janvier 2013*. http://yves.chevallard.free.fr/spip/spip/article.php3?id_article=211.

Cohn, N. (2013). *The visual language of comics: Introduction to the structure and cognition of sequential images*. Bloomsbury Advances in Semiotics.

Cunningham, D. J. (1998). Cognition as semiosis: The role of inference. *Theory and Psychology*, 8, 827-840

De Hosson, C., Bordenave, L. Daures, P.L., Décamp, N., Hache, C., Horoks, J., Kermen, I. (2019). Quand l'élève devient auteur.e : analyse didactique d' ateliers BD-sciences. *Trema*, 51.

Farinella, M. (2018). The potential of comics in science communication. *JCOM* 17 (01), Y01. https://doi.org/10.22323/2.17010401.

Gonzales-Espada, W. J. (2003). Integrating physical science and the graphic arts with scientifically accurate comic strips: rationale, description, and implementation. *Enseñanza de las Ciencias*, 2(1), 58–66.

Lo Iacono, G. and de Paula, A.S.A.T. (2011). A pilot project to encourage scientific debate in schools. Comics written and peer reviewed by young learners. *JCOM* 10(3), A04.

McCloud, S. (1993). *Understanding comics : the invisible art*. (Reprint edition). New York, NY, U.S.A.: William Morrow Paperbacks.

Rabardel, P. (1995). *Les hommes et les technologies ; approche cognitive des instruments contemporains*. Paris : Armand Colin.

Sousanis, N. (2015). *Unflattening*. Harvard University Press.

Tatalovic, M. (2009). Science comics as tools for science education and communication: a brief, exploratory study, *JCOM* 08(04) A02.

Semantic Waves for Revealing Different Levels of Knowledge in Middle-School Science Classes

Anna-Vera Meidell Sigsgaard

University College of Copenhagen, Dept. of Teacher Education, Denmark

This paper reports on an educational research and development project in progress. The project aims to make science lessons in the Danish middle-school (grades 4-6) more accessible to all students by increasing science-teachers' awareness of the linguistic challenges second-language learners in particular (can) encounter in science lessons and materials. The aim of the paper is to demonstrate teacher-development and strategies for making 'science knowledge' visible to all students by integrating meaningful language work in science lessons. The project takes its point of departure in an understanding of students' learning where linguistic interactions are seen as crucial for accessing scientific understandings. In the paper we introduce an analytical approach using semantic waves from Legitimation Code Theory (LCT). Observations and transcriptions of conversations from 4th grade science lessons provide data for semantic gravity analysis, which leads to discussions of how awareness of semantic waves can encourage teachers to integrate language activities in science lessons in order to scaffold their students' science understandings and participation.

Keywords: Middle-school science teaching, Legitimation Code Theory, semantic gravity

Introduction

One of the well-known challenges in science and technology teaching in upper-primary/middle school years is ensuring that students develop understandings of scientific concepts based on the many hands-on activities science lessons often include (Musharrat, 2020). However, participating in experiments and doing hands-on activities does not in itself ensure that students understand the science in question or learn to communicate about it actively (Leach & Scott, 2003; Lemke, 1998).

Language can play a significant role in the difficulties students face, particularly for students who are second-language learners (Gee, 2008; O'Hallaron et al., 2015). One reason for some of these difficulties is that the material covered in middle school science classes is often more complex than what students have learned in previous years. This becomes especially evident in the texts students read and are expected to produce (Christie, 1998; Polias, 2016; Schleppegrell,

2004). International research and educational development indicate that a functional understanding of language (Fang & Schleppegrell, 2014; Lee, 2005; Rose, 2011) together with an explicit focus on the science content (Doran et al., 2021; Macnaught et al., 2013) can offer effective teaching approaches and give more students access to the academic material. Most of these studies, however, are aimed at older students in secondary and tertiary contexts. The overall research question for the project is therefore:

How can analytical and didactic considerations, drawn from respectively LCT (semantic waves) and SFL (linguistic snail), qualify middle school science instruction to the benefit of all students and second-language learners in particular?

The aim of this paper is to demonstrate how semantic gravity analysis can aid in making different kinds of knowledge patterns in pedagogic practice visible to science teachers by introducing them to the concept of *semantic gravity* (Maton, 2013) and the related notion of *semantic waves* (Macnaught et al., 2013; Maton, 2014). By analyzing selected sequences from observed science lessons, three levels of semantic gravity are identified, and teachers' response to these analyses are discussed.

Method

The three-year project reported on is a collaboration involving teacher-educators from Science and Danish as a Second-language units in the Department of Teacher Education, University College Copenhagen (UCC). Inspired by design-based research (Barab & Squire, 2004), this interventionist study had teacher-educators working closely with in-service teachers at four schools in the greater Copenhagen area. Common for all the participants in the project was an interest in developing an explicit language focus in middle-school science lessons to support all students, with a special focus on supporting second-language learners, by planning, teaching and evaluating three units of study at each school.

The participating teachers were introduced to the notions of *semantic gravity* and *semantic waves* from Legitimation Code Theory (LCT) (Georgiou et al., 2014; Maton, 2014) and to the systemic functional linguistics (SFL)-based pedagogic model known as the register model or the snail model (Derewianka, 1990). Through close collaboration between the project's teachers and teacher-educators, a number of science units taught in the school years 2021-22 (at the first school) and 2022-23 (at three additional schools) were (re)designed to increase students' science literacy. Planning and sharing seminars were held throughout the two intervention years, to ensure the teachers from across the three schools had time for sharing ideas, developing language activities and reflecting on their experiences from the project.

Data

The data collected for the project includes interviews with teachers and students, observations from classroom interventions, as well as samples of the used teaching materials and student products. Data collection follows ethics standards and the Danish Code of Conduct for Research Integrity (Ministry of Higher Education and Science, 2014).

For the purposes of this paper, we focus only on data collected from the first school during the first year of interventions (2021-22), including both (video recorded) observations and reflection meetings (audio recorded) with the two involved teachers. Selected sequences from the observed science lessons were transcribed and then analyzed with semantic gravity. These form the point of departure for the following analysis and discussion.

Theory: Semantic Gravity and Semantic Waves for Seeing Different Kinds of Knowledge in Pedagogic Practice

Legitimation Code Theory (LCT) is a sociological framework for researching and informing eg. pedagogic practice (Macnaught, Maton, James R Martin, et al. 2013; Maton 2014). LCT has a close relation with empirical research and offers analytical tools or concepts for analyzing organizing principles underlying social practices. In this paper we focus on only one concept, namely, semantic gravity.

"Semantic Gravity (SG) refers to the degree to which meaning refers to its context" (Maton, 2013, p. 11). It can be understood as the degree of abstraction of a meaning or how closely a meaning is dependent on a context to be understood. Semantic gravity can be conceptualized along a continuum og strengths, from stronger, to weaker (*ibid.*). The *less abstract* a meaning is, i.e. the greater the dependence of a meaning on the context, the *stronger* the semantic gravity (represented by SG+). And vice versa: the *more abstract*, less context-dependent a meaning is, the weaker the semantic gravity (represented by SG-). "All meanings relate to a context of some kind; semantic gravity conceptualizes how much they depend on that context to make sense" (Maton, 2013, p. 11).

Semantic gravity can be used to analyze changes in meanings being made over time in e.g., a science lesson or text, by describing processes of weakening semantic gravity, as when moving from the concrete particulars of a specific science experiment done in a lesson toward abstractions and more general understandings of the scientific notion. Conversely, strengthening semantic gravity occurs when moving from abstract or generalized ideas toward concrete and specific examples or demonstrations of a scientific concept. Movement up and down through various strengths of semantic gravity results in what has been termed semantic waves (Maton, 2013) and can be a useful notion for enacting

changes in pedagogy (Blackie, 2014; Clarence, 2017; Macnaught et al., 2013; Meidell Sigsgaard, 2020).

Inspired by (Georgiou, 2016) we operate with three strengths of semantic gravity: stronger semantic gravity (SG+), weaker semantic gravity (SG-) and in between these, intermediate semantic gravity (SG0) (see Figure 1, below).

Figure 1. Three levels of semantic gravity.

	Coding categories for strengths of semantic gravity	Description of coded content	Examples from our data
semantic gravity	Weaker semantic gravity (SG-)	Meaning expressing a scientific concept or understanding without referencing a specific situation.	A thing can float if it can displace more water than it weighs in itself
	Intermediate semantic gravity (SG0)	Meanings expressing general tendencies and/or explanations	"It was [floated, *ed.*] better with the bigger surface"
	Stronger semantic gravity (SG+)	Descriptions of the conducted experiments/specific contexts in the classroom	"our boat sank when we put too many screws on it"

The three levels of semantic gravity identified in the context of this study are color coded in green, for strongest semantic gravity (SG+): meanings with this strength of semantic gravity are associated with experiences and examples closely dependent on the context of the classroom. When students talk about how their aluminum boat sank after loading it with a certain number of screws, this is a meaning taken directly from the context of their classroom experience, and thus categorized as the strongest level of semantic gravity. When a student compares several different boats' constructions, and based on that generalizes about what makes something float, the meaning this student is making is less directly related to the context of one specific experiment and is thus categorized as intermediate semantic gravity (SG0), here coded with yellow. An excerpt from a text about what makes something float, such as the quote in blue in Figure 1 above, is categorized as the weakest semantic gravity (SG-) as the meaning being made here still concerns floating, but is not at all related to the context of the classroom or even boats. Using these three strengths of semantic gravity allows us to analyze e.g. classroom conversations, to explore what kinds of meanings are being made and by whom (the teacher, students and/or texts/ teaching materials) as demonstrated below in the Findings/Results section.

Findings / Results

Using semantic gravity analysis of observed teacher-student interactions we find that most of the meanings being made by the students are very closely bound to the students' own experiences. By color coding the speaking 'turns' (Sacks

et al., 1974) using green for stronger semantic gravity, yellow for intermediate semantic gravity and blue for weakest semantic gravity we can see which kinds of understandings are being made, and by whom:

Teacher: …Why do you think it could carry so many nails?

Student: We used those popsicle sticks to make the walls. Then the walls were a little stronger, and then we also had a lot of aluminum foil

Teacher: So the shape, the thing with the walls it actually makes a difference

Student: and then there were also popsicle sticks underneath

This exchange is from the end of a day in the science lab, where the fourth-grade class had completed various floating/sinking investigations. Among the day's activities, students worked in pairs to build aluminum foil boats according to different specifications, e.g., without walls, with walls/edges and using straws and/or popsicle sticks made of wood to strengthen their constructions. By loading the boats with screws or nails, the students were to observe which constructions could carry more weight, and based on that, develop an understanding of what is needed for a (heavy) object to float. A whole-class discussion lead by the teacher concluded the day, in which she asked, "What were good techniques for building your boats? Tell me what worked well."

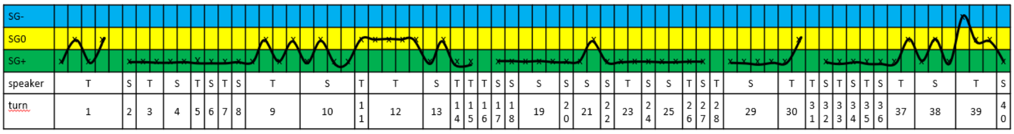
The above excerpt of a teacher-student interaction is from this whole-class discussion at the end of the day, in which the teacher asks the students to generalize based on the day's activities. She is looking for explanations of "what worked well": effectively, she is asking the students to make educated guesses about what makes something float.

The excerpt above between the teacher and a student consists of four turns, two each by the teacher and the student. By coding each turn in terms of the three strengths of semantic gravity, we see that the teacher is making, and asking for, less-context dependent (more abstracted) meanings (i.e. SG0 and SG-, coded in yellow and blue, respectively). The student's responses, however, are mostly very context dependent, referring directly back to the last of their attempts at building an aluminum boat using "a lot of aluminum foil" and "popsicle sticks". The meanings the student is making is therefore coded with green, i.e. stronger semantic gravity.

Semantic gravity analysis of longer passages of exchanges during this whole-class discussion follow this pattern with the teacher asking or rephrasing students' very context-dependent meanings to more generalizable explanations. Figure 2 below shows approximately six minutes from this whole-class discussion, which

lasted about 17 minutes total. The four turns referred to above are the last four turns in Figure 2 below (end of turn 37-40).

Figure 2. Semantic Gravity analysis of approximately 6 minutes of classroom talk.



In Figure 2 above, each turn is numbered (1-40) and shows who is speaking: T for teacher, S for a student. In the analysis, we do not distinguish between students (i.e. all students are coded with S), and do not code turns *not* referring to the scientific concept in question (e.g. when the teacher nominates a student to speak or instructs a student to sit, as in turn 16 and 28).

In a subsequent reflection session with the involved teachers, they were shown a video-clip of the whole-class discussion described above, including the same 6-minute passage of conversation analyzed above. They were then asked to 'analyze', or rather make observations on the video-clip from the perspective of the three strengths of semantic gravity described. The teachers observed that that most of the talk was very strongly anchored in the various experiments the students had made (SG+), and that the students did not make many meanings that were generalized (SG0) and none at all, that were *not* context-dependent (SG-). Another observation made by the teachers was that when the conversation did 'wave upwards' (Macnaught et al., 2013), it was usually done by the teacher, i.e. by rewording the students' responses. An example of this can be seen in the transcription above, when she says, "So the shape, the thing with the walls; it actually makes a difference."

Asking the teachers to pay attention to the meanings being made in a small excerpt of their teaching from the perspective of semantic gravity resulted in what they referred to as an 'Aha' moment. They came to the same conclusions as our analysis here shows. Despite asking for generalizations to sum up the day's activities, students were mainly referring directly to what they had done in their individual groups. As one of the teachers put it, "the conversation isn't really waving." This shows that they realized that they, and not their students, were doing most of the work in bringing the conversation 'upwards' toward more generalized meanings. This observation encouraged the participating teachers to think about what kinds of questions elicit which kinds of answers from the students as demonstrated in a question posed by one of the teachers about her own teaching: "how do we get the kids to wave upwards when they answer?" This indicates that introducing semantic gravity, and the notion of semantic

waves to teachers can be a tool for seeing students' challenges in bridging the divide in meanings between hands-on activities and more abstracted scientific understanding, as well as, potentially, a tool for rethinking or planning lessons to support that transition (Meidell Sigsgaard & Jacobsen, 2021).

Discussion

Results from the project suggest that knowledge of semantic gravity and of semantic waves can be a powerful tool for increasing teachers' awareness (Meidell Sigsgaard et al., 2023) of the epistemology of (Doran et al., 2021; Macnaught et al., 2013; Maton, 2013) science in the middle-grades. We propose that enabling teachers to conceptualize the three levels of context-dependent knowledge in their science lessons, as well as the differences in how students make meanings through language at each of these levels (Meidell Sigsgaard, forthcoming), can enable teachers to focus on integrating meaningful language-work in science lessons, with the aim of explicitly teaching 'science knowledge' to all students.

By asking teachers to think about their teaching from the perspective of semantic gravity, they become empowered to see for example who of their students need help in making meanings that are 'more scientific'. Similarly, teachers started using semantic gravity to look at the teaching materials used in their science (and other subject) lessons. More importantly, understanding the three strengths of semantic gravity presented here, resulted in a desire to include activities in future lessons with the purpose of scaffolding their student's ability to make more generalized meanings, for example by encouraging them to use appropriate scientific vocabulary in their explanations. This enables the teachers to become aware of potential difficulties their students (and second-language learners especially) might encounter in reading such texts. Combined with growing knowledge of a functional model of language (Brisk, 2015; Derewianka & Jones, 2012), this provides ideas for instructional approaches integrating reading and writing activities in science lessons, without losing focus of the content knowledge being taught.

To support second-language learners and all students in understanding and engaging with scientific concepts, it is important to provide explicit language instruction that takes its departure in the scientific concepts being taught in a way that includes teaching scientific literacy (Derewianka, 1990; Gibbons, 2018; Polias, 2016; Windschitl et al., 2018). This includes not only explicitly teaching the specific vocabulary, terms, and conventions of scientific language, but also incorporating these with attention to conventions of scientific writing and communication that are used to convey scientific information.

Acknowledgement

This project has been generously supported by the Novo Nordisk Foundation with grant reference number NNF20OC0065719 in the call "Project grants for natural science education and informal learning environments – Fall 2020".

References

Barab, S., & Squire, K. (2004). Design-Based Research: Putting a Stake in the Ground. *The Journal of the Learning Sciences*, *13*(1), 1–14. https://doi.org/10.1207/s15327809jls1301_1

Blackie, M. A. L. (2014). Creating semantic waves: Using Legitimation Code Theory as a tool to aid the teaching of chemistry. *Chemistry Education Research and Practice*, *15*(4), 462–469. https://doi.org/10.1039/C4RP00147H

Brisk, M. E. (2015). *Engaging Students in Academic Literacies Genre-based Pedagogy for K-5 Classrooms*. Routledge. https://www.routledge.com/Engaging-Students-in-Academic-Literacies-Genre-based-Pedagogy-for-K-5-Classrooms/Brisk/p/book/9780415737043

Christie, F. (1998). Learning the literacies of primary and secondary schooling. In F. Christie & R. Misson (Eds.), *Literacy and Schooling* (pp. 47–73). Routledge. https://doi.org/10.4324/9780203019559

Clarence, S. (2017). Surfing the waves of learning: Enacting a Semantics analysis of teaching in a first-year Law course. *Higher Education Research and Development*, *36*(5). https://doi.org/10.1080/07294360.2016.1263831

Derewianka, B. (1990). Derewianka 1990 Rocks in the Head.pdf. In R. Carter (Ed.), *Knowledge About Language and the Curriculum: The LINC Reader* (pp. 197–215). Hodder and Stoughton.

Derewianka, B., & Jones, P. (2012). *Teaching Language in context*. Oxford University Press.

Doran, Y. J., Maton, K., & Martin, J. R. (2021). The teaching of science: New insights into knowledge, language and pedagogy. *Teaching Science: Knowledge, Language, Pedagogy*, 1–19. https://doi.org/10.4324/9781351129282-1/TEACHING-SCIENCE-DORAN-KARL-MATON-MARTIN

Fang, Z., & Schleppegrell, M. (2014). *Reading in Secondary Content Areas, A language based pedagogy*. Michigan Teacher Training.

Gee, J. P. (2008). *Getting over the slump: Innovation Strategies to Promote Children's Learning*. The Joan Ganz Cooney Center at Sesame Workshop.

Georgiou, H. (2016). Putting physics knowledge in the hot seat: The semantics of student understandings of thermodynamics. In K. Maton, S. Hood, & S. Shay (Eds.), *Knowledge-building: Educational studies in Legitimation Code Theory* (pp. 176–192). Routledge. https://doi.org/10.4324/9781315672342-19

Georgiou, H., Maton, K., & Sharma, M. (2014). Recovering Knowledge for Science Education Research: Exploring the "Icarus Effect" in Student Work. *Canadian Journal of Science, Mathematics and Technology Education*, *14*(3), 252–268.

https://doi.org/10.1080/14926156.2014.935526

Gibbons, P. (2018). *Bridging Discourses in the ESL Classroom: Students, Teachers and Researchers*. Bloomsbury.

Leach, J., & Scott, P. (2003). Individual and Sociocultural Views of Learning in Science Education. *Science & Education*, *12*(1), 91–113. https://doi.org/10.1023/A:1022665519862

Lee, O. (2005). Science Education With English Language Learners: Synthesis and Research Agenda. *Review of Educational Research*, *75*(4), 491–530. https://doi.org/10.3102/00346543075004491

Lemke, J. L. (1998). Multiplying meaning: Visual and verbal semiotics in scientific text. In J. R. Martin & R. Veel (Eds.), *Reading Science* (pp. 87–113). Routledge. https://www.researchgate.net/publication/246905867

Macnaught, L., Maton, K., Martin, J. R., & Matruglio, E. (2013). Jointly Constructing Semantic Waves: Implications for teacher training. *Linguistics and Education*, *24*, 50–63. https://doi.org/10.1016/j.linged.2012.11.008

Maton, K. (2013). Making semantic waves: A key to cumulative knowledge-building. *Linguistics and Education*, *24*(1), 8–22. https://doi.org/10.1016/j.linged.2012.11.005

Maton, K. (2014). Knowledge and knowers: Towards a realist sociology of education. In *Knowledge and Knowers: Towards a realist sociology of education*. Routledge. https://doi.org/10.4324/9780203885734

Meidell Sigsgaard, A.-V. (forthcoming). Supporting upper primary students' science understandings through dialogic interactions and language support activities (working title). In Fitzgerald, Ange, Earle, Sarah, Preston, Chris, & Georgiou , Helen (Eds.), *Primary Science Learning for Children, Teachers, and Communities—Stories of Practice and Possibility for Science Educators*. Springer.

Meidell Sigsgaard, A.-V. (2020). Making waves in teacher education: Scaffolding in 'teaching bilingual students'. In C. Winberg, S. McKenna, & K. Wilmot (Eds.), *Building Knowledge in Higher Education: Enhancing teaching and learning with Legitimation Code Theory*. Routledge.

Meidell Sigsgaard, A.-V., Heinrich, S., Pagaard, D. M., & Olsen, P. S. (2023). Semantiske Bølger: Sprogligt arbejde i natur/ teknologiundervisningen med fagligt fokus. *MONA*, *3*, 26–43.

Meidell Sigsgaard, A.-V., & Jacobsen, S. K. (2021). Semantiske bølger: Analytisk redskab og didaktisk princip. In S. K. Knudsen & L. Wulff (Eds.), *Kom ind i sproget: Flersprogede elever i fagundervisningen* (pp. 73–90). Akademisk Forlag.

Ministry of Higher Education and Science. (2014). *The Danish Code of Conduct for Research Integrity*. https://ufm.dk/en/publications/2014/the-danish-code-of-conduct-for-research-integrity

Musharrat, T. (2020). Teachers' Perceptions About Use and Challenges of Hands-On Activities in Secondary Science Classrooms. *European Journal of Education*

Studies, *7*(12). https://doi.org/10.46827/ejes.v7i12.3384

O'Hallaron, C. L., Palincsar, A. S., & Schleppegrell, M. J. (2015). Reading science: Using systemic functional linguistics to support critical language awareness. *Linguistics and Education*, *32*. https://doi.org/10.1016/j.linged.2015.02.002

Polias, J. (2016). *Apprenticing Students into Science—Doing, talking & writing scientifically*. Lexis Education.

Rose, D. (2011). Beating educational inequality with an integrated reading pedagogy. In F. Christie & A. Simpson (Eds.), *Literacy and Social Responsibility: Multiple Perspectives* (pp. 101–115). Equinox.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, *50*(4), 696–735. https://doi.org/10.2307/412243

Schleppegrell, M. J. (2004). *The Language of Schooling: A functional linguistics perspective*. Erlbaum.

Windschitl, M., Thompson, J. J., & Braaten, M. L. (2018). *Ambitious science teaching*. Harvard University Press.

Digital Research Skills: Application in Secondary Science Education

Kim Blankendaal[1], R. F. G. Meulenbroek[1], W. R. Van Joolingen[1]

Utrecht University, Utrecht, Netherlands

This study focuses on the current level of Digital Research Skills (DRS) as applied by pre-university science students. A rubric containing 6 categories and 38 categories was constructed and tested based on a DRS framework. 88 randomly collected students' science project examination reports were analysed using the rubric. The results show that students demonstrate difficulties in analysing, transforming and visualizing content/data digitally and writing a research paper using digital tools. For example, graphs in the reports frequently only contain a line without measurement points, while axis values are displayed inconsistently or are missing. Furthermore, formulas are often visualized without a proper format. To bridge the gap in the transition to an academic science study, more attention on DRS in secondary education is warranted.

Keywords: Digital research skills, secondary education, science project report.

Introduction

A major goal of secondary science education is preparing students for an academic career in science and technology. This implies the development of skills needed in (learning to) doing research, especially for those using digital technology. This comprises skills for collecting, handling and presenting data and sources to write a research paper. We label these skills *Digital Research Skills* (DRS) (Author, 2023).

First-year university students are confronted with an environment that often draws upon such skills. For example, Akuegwu and Uche (2019) found that while reading, presentation, communication and information-gathering skills of freshman STEM students were adequate, especially data analysis was found wanting. Other research confirmed this problematic level of skills such as information seeking and using ICT among students in secondary education (Udeogalanya, 2022; Walraven et al., 2008)institutions of higher education transitioned to fully online teaching and learning. Prior to the pandemic, fully online education was secondary to face-to-face format. Only about 20% of classes were fully online while 80% were face-to-face. Digital literacy was given a

token treatment as a percentage of the entire curriculum and relegated to only the departments of computer information systems and computer sciences. Faculty, students, staff, families, and communities were not trained for the intensity of fully online education as the only format. Many of them never heard of most of the digital literacy tools. Both faculty and students were forced to learn the use of computers and digital literacy tools to survive the spring 2020 semester. Low-income families were left to the operational schedules of local libraries, many of whom did not have internet presence. The institutions provided limited training for faculty and students to meet the urgency of the time. Faculty and students were forced to purchase expensive tools and hardware to withstand the intensive demand of teaching and learning. Many students were overwhelmed by the pressure of the new way of learning; some dropped out of school while many performed very poorly in their examinations which negatively impacted their overall grade point averages. One year later in spring 2021, the student success outcomes have barely changed. At the same time, technology is advancing despite the raging COVID pandemic. Millions of technology-enabled jobs remain unfilled while millions of university graduates are unemployed. There continues to be a mismatch between current job requirements in the industries and graduating students' skills. This paper discusses the indispensable value of building computer and digital literacy training into all undergraduate curriculums. We argue for mandated computer and digital literacy exit skills assessment …","author":[{"dropping-particle":"","family":"Udeogalanya","given":"Vero nica","non-dropping-particle":"","parse-names":false,"suffix":""}],"container-title":"International Journal of Higher Education Management","id":"ITEM-1","issue":"2","issued":{"date-parts":[["2022","2"]]},"language":"English","n ote":"Copyright - © 2022. This work is licensed under https://creativecommons. org/licenses/by/3.0/ (the "License".

Unfortunately, since it is often unclear at what level DRS are addressed in secondary science education (Voogt et al., 2013), it is also largely unknown to what extent students need more practice to prepare them for the transition to a tertiary science study. The main aim of this study is to investigate to what level Digital Research Skills (DRS) are applied by students in secondary science education, by analysing the level of DRS of pre-university science students in their final science project reports. Our research question is: What is the current level of digital research skills (DRS) of pre-university students as evidenced in their research reports?

Method

Based on a framework for DRS (Author, 2023) we constructed a rubric to analyse students' DRS as expressed in research project reports. We applied this to Science Project Reports (SPR) students wrote as part of their examination.

The rubric contains 6 categories and 38 subcategories. Most subcategories appear in DRS category Analyse, transform and visualize content/data digitally' and category 'Write a research paper using digital tools.'

'Analysing, transforming and visualizing content/data digitally' consists of four subcategories: (a) digitally processing and analysing data, (b) digitally transforming and formatting data in graphs, (c) digitally transforming and formatting data in tables, and (d) using other data, such as images, reaction equations, and programming codes. 'Writing a research paper using digital tools' consists of five subcategories: (a) use of automatic table of content and numbering of e.g. pages, figures and tables, (b) use of the insertion caption function for e.g. figures and tables, (c) correctly digital formatting of text, e.g., the use of subscript/superscript, italics, styles and headings, symbols, bullet points, line spacing and margins, (d) digital drawing of the research design or technical design and (e) correct use of a formula editor.

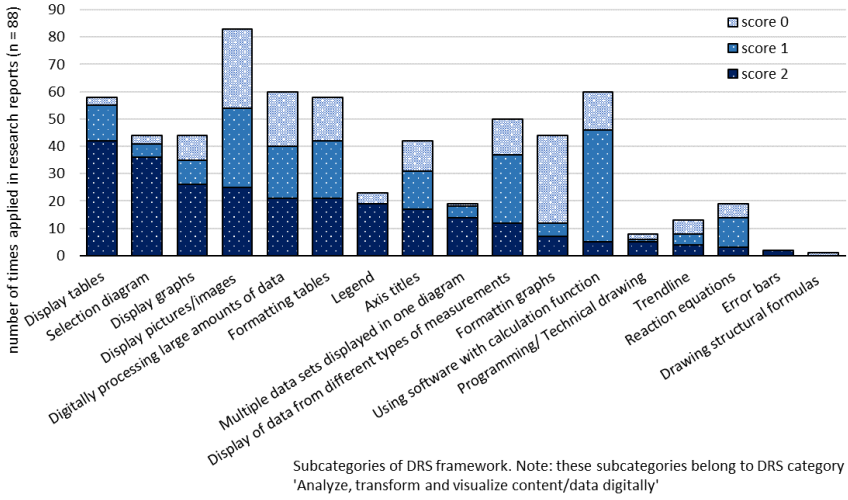
All subcategories are assessed on three levels (score 0, 1 and 2), where 2 means correct application of a DRS, 1 a partial correct application and 0 an incorrect or absent application. Using dual coders, the rubric was fine-tuned in a pilot.

We collected eighty-eight pseudonymized SPRs of randomly approached graduated pre-university science students in the Netherlands to perform a document analysis. The collected reports describe research in physics (n = 24), chemistry (n = 48) and nature, life and technology (NLT) (n = 16), the latter being an interdisciplinary advanced science topic in the Dutch curriculum.

Results

In Figure 1, the obtained scores in DRS category 'Analyse, transform and visualize content/data digitally,' are shown. The subcategories are ordered descending based on number of assigned score 2.

Figure 1: Application of Digital Research Skills (DRS) in students' science reports (N = 88). Score 2 = correct, score 1 = partial correct and score 0 = incorrect/absent.



Zooming in on the skills in DRS category 'Analyse, transform and visualize content/data digitally,' we see that the majority of the SPRs contains collected quantitative data (60 out of 88). In 44 out of 88 reports the data are presented in a chart, with graphs that are displayed correctly (score 2; 26 of 44). A part is displayed almost correctly (score 1; 9 of 44), for example, by displaying axis values inconsistently or not using secondary grid lines, making the measurement points difficult to read. The remaining SPRs scored 0 points (9 of 44), for example, by only drawing a line without measurement points. Of the 44 research reports where a graph was displayed, 37 reports did not score points for formatting the graph. For example, because the graph contained a title above it (score 0) or because of a dark grey outline around the graph (score 1). These two elements are added by default in Excel and were not removed by the students. Remarkably, in 14 out of 42 SPRs that contain a graph, one of the axis titles is incorrect (score 1) or missing (score 0; 11 of 42).

Furthermore, it can also be seen that in 13 of the 58 research papers involving an experiment, the student added a trend line. A portion did not have the trend line drawn through the origin when it should have or did not update the trend line formula with the correct quantities or units (score 1, 4 of 13). On the other hand, a line was drawn from measurement point to measurement point instead of a line through the measurement points (score 0; 5 of 13).

The obtained scores in DRS category 'Write a research paper using digital tools,' show a similar pattern to Figure 1. Major shortcomings in 14 of the 16 subcategories are identified. For example, inserted figures and tables do not always contain correct numbering (score 1 and 0; 57 of 86) and a caption (score

1 and 0; 49 of 88). Furthermore, physical quantities and chemical formulas are very often used without applying italics (score 1 and 0; 30 of 38) or sub- and superscripts (score 1 and 0; 30 of 60). Students make little use of the formula editor in the word processor. Only 5 research reports show correct use of the formula editor. The majority show an inability to use the formula editor correctly (score 1; 22 of 40). It is problematical that a certain number of students do not use the formula editor at all (score 0, 13 of 40).

Conclusion and Discussion

Our results show that secondary science students are able to apply multiple DRS in a Science Project Report (SPR). However, the current level of DRS, as visible in research reports of graduating pre-university science students, is lacking in certain areas. Students are generally unable to properly use software with calculation function, e.g., to process and display data from different types of measurements, to draw and format a graph, to make use of a trend line and correct axis titles. 'Writing a research paper using digital tools' shows a lack of skill in certain areas as well. We found that students lack skills to format a graph or table, use automatic numbering and add a caption. Furthermore, the use of a formula editor, including italicization of quantities and upright units also needs attention, and last but not least the level of formatting a text is also inadequate.

The limitation of this study is mainly that the sample of SPRs was exclusively from Dutch students. However, the general needs cited in the introduction showed that the issues addressed are encountered in other countries as well.

To bridge the gap in the transition to an academic science study we recommend more attention and practices on DRS in secondary education. Secondary science students should focus on the application of the following skills: (1) inserting a high quality image, graph or table using automatic numbering and adding a caption (2) the use of software with calculation function to process a display data into a proper graph, including axis titles, trend lines and a sufficient format, (3) the use of a formula editor, including italicization of quantities and upright units, (4) the use of formatting options in a text processor, including styles and headings, sub- and superscripts, text wrapping, margins and alignments.

The DRS rubric can be used as a measuring tool for the development of DRS during secondary education.

References

Akuegwu, B. A., & Uche, K. (2019). Assessing graduate students' acquisition of research skills in universities in Cross River State Nigeria for development of the total person. *Educational Review: International Journal*, 6(1), 27–42.

Blankendaal-Tran, K., Meulenbroeks, R. F. G., & van Joolingen, W. R. (Accepted 2023). Digital Research Skills in secondary science education: A guiding framework and qualitative evaluation. *European Journal of STEM Education*.

Udeogalanya, V. (2022). Aligning digital literacy and student academic success: Lessons learned from COVID-19 pandemic. *International Journal of Higher Education Management*, 8(2). https://doi.org/https://doi.org/10.24052/IJHEM/V08N02/ART-4

Voogt, J., Erstad, O., Dede, C., & Mishra, P. (2013). Challenges to learning and schooling in the digital networked world of the 21st century. *Journal of Computer Assisted Learning*, 29(5), 403–413. https://doi.org/10.1111/jcal.12029

Walraven, A., Brand-Gruwel, S., & Boshuizen, H. P. A. (2008). Information-problem solving: A review of problems students encounter and instructional solutions. *Computers in Human Behavior*, 24(3), 623–648. https://doi.org/10.1016/j.chb.2007.01.030

Evaluating A Teaching-Learning Sequence (TLS) About Acid-Base Reactions In Upper Secondary School

Rita Krebs[1], Marvin Rost[1], Elisabeth Hofer[2], Anja Lembens[1]

[1]Austrian Educational Competence Centre Chemistry, University of Vienna, Austria

[2]Leuphana University Lüneburg, Lüneburg, Germany

Acid-base chemistry plays an important role in our lives because of its real-world applications (biochemical and industrial processes including acid-base reactions, everyday products containing or consisting of acidic or basic compounds, ...). Consequentially, the topic is an important one in chemistry (and science) education. We adapt the Brønsted-Lowry concept, for the upper secondary level accordingly in the Design-Based Research project presented here. The focus is on the compatibility of acid-base reactions to other reaction types (e.g. redox reactions) by using visualizations such as the Electron Pushing Formalism. By means of Educational Reconstruction, five key ideas (KIs) are identified as the central aspects of the topic for the target group. An explanatory framework based on the KIs is presented to upper secondary school students to evaluate acceptance, plausibility, and applicability in three rounds of interviews employing the method of probing acceptance (N_1=7, N_2=4, N_3=7). Based on the results of these assessments, a teaching-learning sequence (TLS) is designed and evaluated in an intervention study (N=57) using two questionnaires: a) A closed-ended Rasch-scaled acid-base knowledge test and b) an open-ended test in an intervention study. The findings from the open-ended questions show a significant increase in test scores obtained via evaluative qualitative text analysis which aligns well with the closed-ended items. Additionally, we observe a shift towards responses that center around the acid-base concept presented in our TLS when conducting a thematic analysis of the learners' answers. We conclude that our intervention is conducive to learning about defining acid-base reactions and therefore recommend employing the Electron Pushing Formalism to teach this definition.

Keywords: Secondary education, acid-base chemistry, design-based research.

Introduction

Acid-base reactions hold significance in our everyday lives and routines, which is, for example, evident in the fact that related scientific terms were integrated into everyday language (Bučková & Prokša, 2021). *Neutralization* and *oxygen*, terms

historically associated with acid-base reactions and the development of different acid-base concepts, have been successfully consolidated into everyday speech. The different acid-base concepts that led to the coinage of these terms have been developed over centuries, leading to a plethora of different explanations of what constitutes and acid and a base (Furió-Más et al., 2005; Jiménez-Liso et al., 2020; Kousathana et al., 2005; Krebs & Hofer, 2022). Although these concepts and models share a common goal of modeling acid-base reactions and defining 'acids' and 'bases', they diverge in the levels of the Johnstone triangle they attach to (Taber, 2013). For example, Boyle's (1675) acid-base concept only refers to macroscopic properties of acidic and alkaline solutions, i.e., changing the colour of plant dyes such as litmus (Kousathana et al., 2005), whereas Lewis' (1923) concept explains the behavior of the substances by referring to underlying reaction mechanisms. Whilst the numerous different acid-base concepts afford us a lens through which we can comprehend how chemical classifications, concepts, and models have evolved over time (Erduran & Kaya, 2019; Jiménez-Liso et al., 2020), this variety of concepts may also lead to (learner) confusion, conceptual confusion, and the development of undesirable alternative and hybrid conceptions (Alvarado et al., 2015; Hoe & Subramaniam, 2016; Justi & Gilbert, 1999, 2000; Lembens et al., 2019; Pan & Henriques, 2015). The plethora of acid-base models has led to the design of numerous instructional approaches and teaching-learning sequences about acid-base chemistry (e.g. Erduran & Kaya, 2019; Jiménez-Liso et al., 2020; Ültay & Calik, 2016; Watson et al., 2020). However, little attention has been devoted to two of the major obstacles connected with acid-base reactions in chemistry lessons: Firstly, the idea that particles termed 'acids' and 'bases' offer the same properties as the associated substances and secondly, the lack of compatibility to other reaction types such as redox reactions and reaction mechanisms in organic chemistry. The former has, for example, been observed when teaching about matter in that learners fail to understand properties of a bulk matter as emergence (Albanese & Vicentini, 1997; Budimeier & Hopf, 2022; Derman et al., 2019). The latter may, for instance, produce alternative conceptions such as the idea that all "[s]pecies having formulas with hydrogen are acids", as the underlying process of what constitutes an acid (particle) is not made evident (Demircioglu et al., 2005, p. 46). These challenges, in turn, reinforce de Vos' and Pilots' (2001, p. 494) position towards acid-base reactions in chemistry as a "sedimentary rock". They explain that the topic has "acquired a complicated conceptual structure" as, apparently, "it was much easier to add something to the existing curriculum than to remove something from it or to restructure it" (de Vos & Pilot, 2001, p. 494).

We address their call for restructuring explanatory frameworks for acid-base reactions as follows: First, designing a teaching-learning sequence (TLS) that presents a comprehensive concept of acid-base reactions in coherence with

other types of chemical reactions, such as redox and organic reactions. Second, considering the distinction between particle and macroscopic level during the whole sequence. In order to understand students' learning gains from this TLS, we aim at answering the following research questions:

1. What are the main themes identified in the learners' declarative knowledge about acid-base chemistry?
2. What is the effect of our TLS on learners' declarative knowledge about acid-base chemistry?

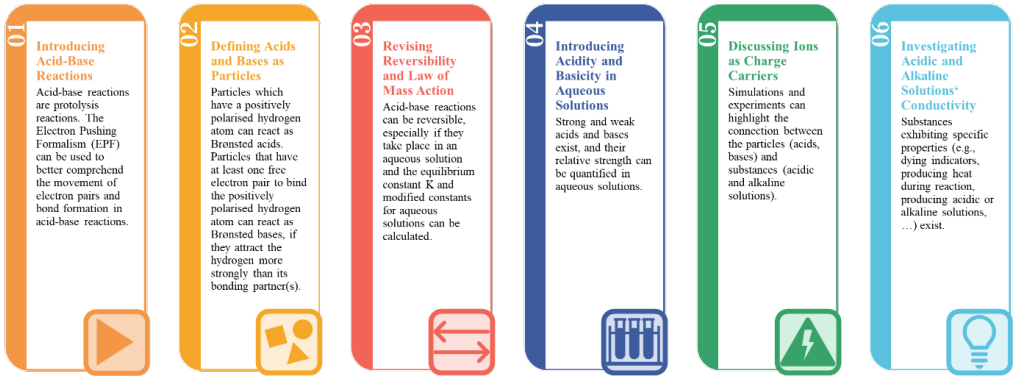
Developing A Teaching-Learning Sequence About Acid-Base Reactions for Upper Secondary School Students

To address the challenges of attributing properties of substances to particles and vice versa as well as a lack of compatibility when teaching about acid-base reactions, the TLS encompasses to following key ideas:

- Introducing acid-base reactions through the electron pushing formalism (EPF) and the utilization of Lewis formulae to highlight similarities and differences to other reaction mechanisms (Ghosh & Berg, 2014; Sieve & Bittorf, 2016).
- Presenting the properties of acids and bases as emerging properties within their specific chemical reactions (Krebs et al., 2022).
- Highlighting the reversibility inherent in acid-base reactions (Jiménez-Liso et al., 2020).
- Explaining the strength of acids and bases by employing beaker models and referencing the pK_a table (Watson et al., 2020).
- Establishing a connection between the submicroscopic and the macroscopic level through experiments and a simulation (Krebs, 2023; Lancaster et al., 2021).

Overall, our emphasis lies in elucidating the reaction mechanism and highlighting electron transfer as a fundamental principle underlying all chemical reactions. In addition, assigning the labels 'acids' and 'bases' to particles during an acid-base reaction is an emerging consequence of their interplay: 'being' an acid or a base is not an intrinsic characteristic of a particle, akin to sulphur atoms not being yellow (Albanese & Vicentini, 1997; Barke, 2006). On the contrary, the bulk properties of an alkaline or acidic solution or gas only exist because of reactions occurring between particles termed as 'acids' and 'bases'.

Figure 1. The structure of the proposed teaching-learning sequence.

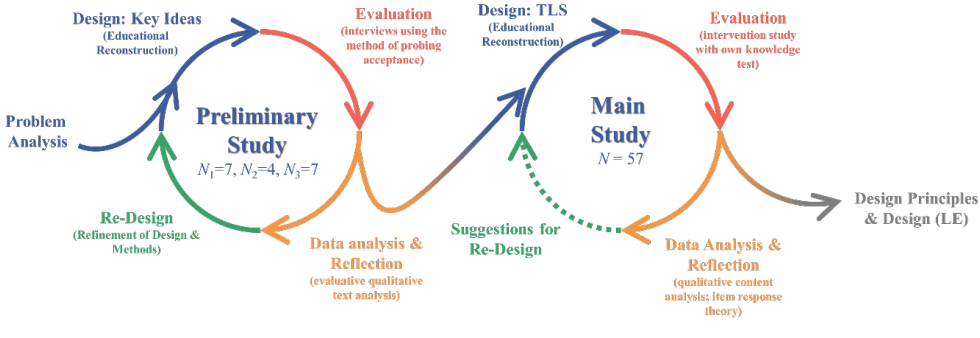


The resulting TLS consists of six units (cf. Fig 1) discussing these key ideas.

Research Design and Methods

As presented in Figure 2, the TLS evaluated during the project emerged from an Educational Reconstruction (Duit et al., 2012) of the topic (Krebs et al., 2022). Based on an analysis of the science content, and research on teaching and learning about acid-base chemistry (Jiménez-Liso et al., 2020; Lembens et al., 2019), the key aspects mentioned above, an explanatory framework and, finally, a larger TLS were constructed (Krebs et al., 2022; Krebs & Lembens, 2021). These key ideas cover the main aspects of the topic regarding the target audience, i.e., local upper secondary learners of chemistry.

Figure 2. The study design, based on the Design-Based Research Cycle, © Sarah Zloklikovits, CC BY-SA 4.0, https://creativecommons.org/licenses/by-sa/4.0, via Wikimedia Commons.



Whereas the explanatory framework as a first step towards a TLS was investigated in the preliminary study (Krebs & Lembens, 2024), the TLS itself was investigated in the form of an intervention study. The intervention study's sample comprised five upper secondary chemistry classes in grades 10 and 11 from two local grammar schools. These grammar schools emphasize languages and natural sciences, providing both standard chemistry classes and laboratory sessions. The primary inclusion criterion for analysis was the availability of two datasets (pre- and post-tests) per student, aligning with their identification codes. Throughout the intervention study, a significant dropout rate was observed ($N_{original}$ = 88, N_{sample} = 57). The students' ages range from 14 to 18, and there were 22 female, 31 male and 2 diverse students, as well as 2 students whose gender was not specified.

We developed and piloted (N_{pilot} = 136) an acid-base knowledge test consisting of 22 closed-ended items and six open-ended questions (Krebs et al., 2023). A Rasch model was calculated to assess item difficulty (Wilson, 2004; Wu et al., 2016), and the 22 items were accordingly divided into eleven pre-test and eleven post-test items. In addition to the closed-ended items, the six open-ended questions were administered to participating learners (N = 57) in pre- and post-test. The learners were prompted to summarize the main idea of the previous chemistry lessons (OF1), and to assess learners' knowledge in defining acid-base reactions (OF2), the terms acid (OF3) and base (OF4) and acidity (OF5) and basicity (OF6).

Table 1. Coding guideline for the evaluative qualitative text analysis with examples for the categories.

Category Name	Category Points	Description	Example
Appropriate definition according to Brønsted-Lowry	5	This category includes all explanations in which the task was solved correctly, and all the important content and central concepts were addressed.	"The particle in an acid-base reaction that releases its hydrogen particle." *OF03, post-test, gender not specified, 15 years old*
Aspects of the Brønsted-Lowry Definition	4	This category includes all explanations in which the task was solved correctly or partially correctly and some important content and central concepts were addressed.	"The hydrogen molecule migrates from the acid to the base." *OF02, post-test, male, 16 years old*
Other Acid-Base Concepts	3	This category includes all explanations in which the task was solved using acid-base concepts other than the desired one.	"Compounds that are capable of forming hydroxide ions in aqueous solutions." *(OF04, post-test, female, 15 years old)*
Paraphrase of Task Description	2	This category includes all statements that are a (partial) paraphrase of the task description.	"A reaction between an acid and a base." *OF02, pre-test, female, 15 years old*
Inadequate Response	1	This category includes all explanations in which the task was not solved appropriately or in which central concepts and theories were used incorrectly.	"A base is a liquid with a pH value below 7." *OF04, pre-test, female, 17 years old*
Missing	0	This category includes fields intentionally left blank as well as statements as to why the task could not be completed.	-

The main themes in the answers to the open-ended questions OF2 to OF6 were identified via thematic analysis (Braun & Clarke, 2022). Themes were identified iteratively until reaching saturation. The themes were then used to design a metric for evaluative qualitative text analysis (Kuckartz, 2014) and assessed with descriptive statistics.

Results

The findings presented in this paper are twofold; first, we report on the findings from the qualitative aspect of the analysis followed by a statistical comparison of the scores gained from the evaluative qualitative text analysis.

Themes Identified In The Students' Answers

The analysis produced five themes, and the additional category *missing* for missing answers.

Inadequate Response

In the pre-test, numerous inadequate answers were given to the open-ended questions. These include incorrect use of scientific terms and allusions to known alternative conceptions. Learners remember basic terms connected to acid-base chemistry but use them in an inaccurate manner: "A base has a very low pH value" *(OF04, pre-test, female, 16 years old).*

Paraphrase Of Task Description

Similarly, many learners paraphrase the task description in their response. According to a 15-year-old male student, acid strength "[p]robably [refers to] the strength of the acid" *(OF05, pre-test)* and an acid-base reaction is defined as "[a] reaction between acid and base" *(OF02, pre-test, female, 15 years old).*

Other Acid-Base Concepts

Another theme identified is the use of acid-base concepts other than the Brønsted-Lowry one. For example, one student uses the Arrhenius (1884) definition of a base to define it: "Compounds that are capable of forming hydroxide ions in aqueous solutions" *(OF04, post-test, female, 15 years old)*. Another uses this concept to define acid-base reactions: "In an acid-base reaction, an acid reacts with a base to form a salt and a water [while releasing] heat (exoterm) [sic]" *(OF02, pre-test, male, 16 years old)*.

Aspects of The Brønsted-Lowry Definition

This theme was identified and coded when important content and central concepts were partially addressed but other aspects were lacking. For example, a 17-year-old male student defined acids as "[p]articles where a positive polarized H atom is present" *(OF3, post-test)*.

Appropriate Definition According To Brønsted-Lowry

Finally, several learners' answers included the Brønsted-Lowry concept appropriately. For example, one student defined an acid-base reaction thus:

> An acid-base reaction is, as the name suggests, a reaction between an acid and a base. The acid, which consists of an H ion and an acid residue, releases the H^+ ions and the base, which is a compound that has a free pair of electrons, absorbs the H ions from the acid. *(OF2, post-test, male, 17 years old)*

The main aspects of the underlying scientific concept are mentioned adequately in these answers.

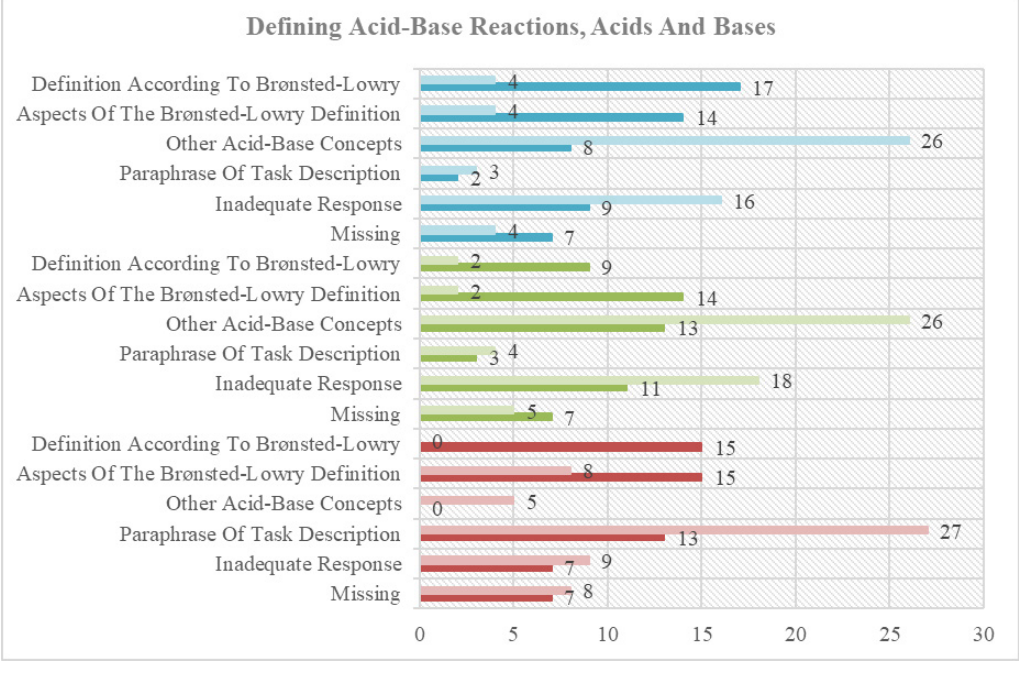
Comparison Of Students' Answers Before and After Completing The TLS

To assess the effect of our TLS on learners' declarative knowledge about acid-base chemistry, we compared both the qualitative change in learners' answers and their average scores in pre- and post-test. For that reason, we used the themes identified in the learners' answers to develop a scoring rubric (Table 1) and assigned points accordingly.

Qualitative Changes In Learners' Answers After Completing The TLS

We identified a shift in the learners' answers from pre- to post-test; the responses changed from offering a paraphrase of the task description when defining acid-base reactions or using alternate acid-base concepts to define the terms 'acid' and 'base' towards including more aspects of the Brønsted-Lowry concept (Figure 3).

Figure 3. The number of categorized learners' (*N*=57) definitions of acid-base reactions (blue) as well as of acids (green) and bases (red) as particles, with light colors corresponding to answers given in the pre-test.

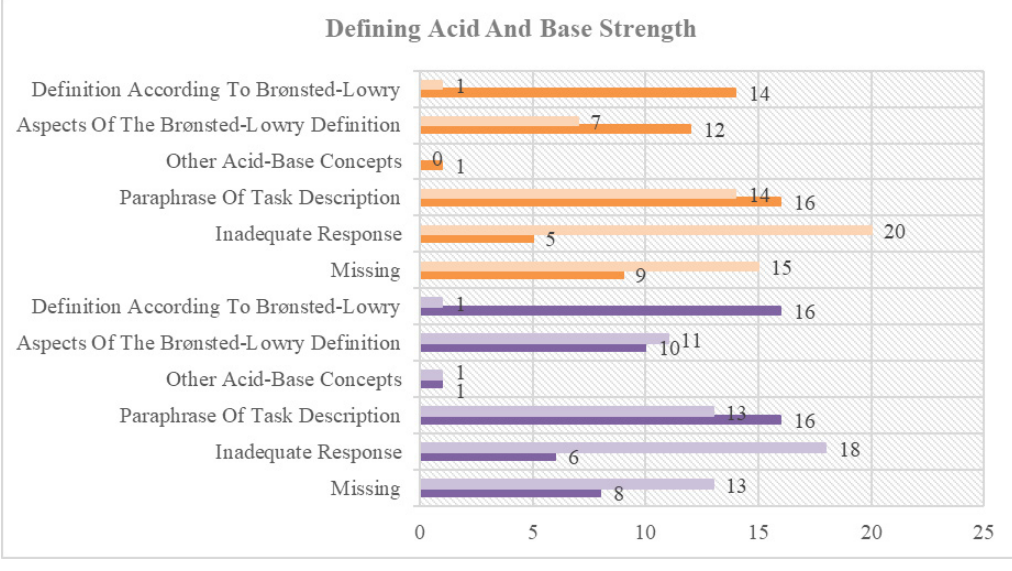


Similarly, the definitions of acid/base strength given in the pre-test aligned with the task description, whereas a quarter of the learners gave a definition aligned the desired concept in the post-test (cf. Figure 4).

Comparing The Learners' Scores Before And After Completing The TLS

To assess the change in learners' scores on the open-ended questions over the course of the intervention, we calculated average scores for the pre- and post-test. Test scores on the open-ended questions were significantly higher after completing the TLS, $t(56) = 4.51$, $p = <.001$, $d = .60$, which is classified as medium efficiency according to Cohen (1988). According to the Rasch model calculated for the multiple-choice items, learners' person ability estimates also increased significantly from pre- to post-test (Krebs et al., 2023).

Figure 4. The number of categorized learners' (*N*=57) definitions of acid (orange) and base strength (purple) as particles, with light colors corresponding to answers given in the pre-test.



Discussion

The main goal of this study was to identify themes in the learners' responses and to assess the effect of our TLS on learners' declarative knowledge about acid-base chemistry. The findings suggest an increase in knowledge both qualitatively and in scores after completing the sequence. For example, in the pre-test 16 learners defined an acid-base reaction inadequately and referred to alternative conceptions known in the literature (Hoe & Subramaniam, 2016; Pan & Henriques, 2015) and less than ten percent of the overall group offered an appropriate definition according to our rubric. To be specific, an 18-year-old male student writes in the pre-test that an acid-base reaction is a "reaction that contains an acid or base and reacts with another substance can cause effects such as dissolution" (OF02), which is similar to the idea that "[a]n acid is something which eats material away or which can burn you" (Hand & Treagust, 1988, p. 55).

After learning about the reaction mechanism of acid-base reactions via Electron Pushing Formalism (Ghosh & Berg, 2014; Sieve & Bittorf, 2016), more than half of the learners offered an adequate or partially adequate definition. For instance, that same student defines an acid-base reaction as a "reaction between an acid and a base according to Bronstein [sic!] with donating and accepting H ions" in the post-test. After exposure to our explanatory framework and beaker models visualizing acid and base strength (Barke, 2006), a similar increase in learners' answers was observed. Moreover, the learners' scores on their answers in the post-test increased significantly in comparison to the post-test. This supports the hypothesis that the TLS induced an increase in declarative knowledge about acid-base reactions that align with the explanatory framework we developed to an extent.

Conclusion and Outlook

In conclusion, our study aimed to evaluate the impact of the TLS on learners' declarative knowledge of acid-base chemistry. The TLS addresses two main challenges: the alternative conception that 'acids' and 'bases' as particles possess identical properties to the associated substances, and the lack of compatibility with other reaction types. The TLS aims to address these challenges in the form of a corresponding explanatory framework and the use of the Electron Pushing Formalism and beaker models as visualizations (Barke, 2006; Ghosh & Berg, 2014; Sieve & Bittorf, 2016). Our findings underscore the efficacy of the TLS to teaching about acid-base chemistry in upper secondary school. We are justified in concluding that restructuring the eclectic field of explanatory frameworks is possible and can be implemented in actual chemistry classes. To develop a full picture of the TLS and its support of developing a compatible understanding of acid-base chemistry, however, research is needed to investigate retention of declarative knowledge as well as knowledge transfer to different reaction types. We propose and pursue a stronger experimental approach to compare our TLS to established learning sequences. Finally, applying the Electron Pushing Formalism in teaching and learning chemistry has the potential for more consistency in cherished but rather historical ideas about chemistry learning and thus contributes to a necessary overarching discussion within our field.

References

Albanese, A., & Vicentini, M. (1997). Why do we believe that an atom is colourless? Reflections about the teaching of the particle model. *Science & Education*, *6*(3), 251–261. https://doi.org/10.1023/A:1017933500475

Alvarado, C., Cañada, F., Garritz, A., & Mellado, V. (2015). Canonical pedagogical content knowledge by CoRes for teaching acid–base chemistry at high school. *Chem. Educ. Res. Pract.*, *16*(3), 603–618. https://doi.org/10.1039/C4RP00125G

Arrhenius, S. (1884). *Recherches Sur La Conductibilite Galvanique Des Electrolytes: Part 1-2.* Kessinger Publishing.

Barke, H.-D. (Ed.). (2006). *Chemiedidaktik: Diagnose und Korrektur von Schülervorstellungen* (1. Aufl.). Springer.

Bianchini, J. A. (2017). Equity in Science Education. In K. S. Taber & B. Akpan (Eds.), *Science Education* (pp. 455–464). SensePublishers. https://doi.org/10.1007/978-94-6300-749-8_33

Boyle, R. (1675). Reflections Upon the Hypothesis of Alcali and Acidum. In R. Boyle (Ed.), *Experiments, notes, &c., about the mechanical origine or production of divers particular qualities: Among which is inserted a discourse of the imperfection of the chymist's doctrine of qualities; together with some reflections upon the hypothesis of alcali and acidum.* E. Flesher.

Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. SAGE.

Bučková, A., & Prokša, M. (2021). The persistence of primary school students' initial ideas about acids and bases in the mental models of adults. *Chem. Educ. Res. Pract.*, *22*(1), 164–174. https://doi.org/10.1039/d0rp00156b

Budimeier, F., & Hopf, M. (2022). STUDENTS' IDEAS ON COMMON EXPERIMENTS ABOUT THE PARTICULATE NATURE OF MATTER. *Journal of Baltic Science Education*, *21*(3), 381–397. https://doi.org/10.33225/jbse/22.21.381

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). N.J: L. Erlbaum Associates.

Demircioglu, G., Ayas, A., & Demircioglu, H. (2005). Conceptual change achieved through a new teaching program on acids and bases. *Chemistry Education Research and Practice*, *6*(1), 36–51. https://doi.org/10.1039/B4RP90003K

Derman, A., Koçak, N., & Eilks, I. (2019). Insights into Components of Prospective Science Teachers' Mental Models and Their Preferred Visual Representations of Atoms. *Education Sciences*, *9*(2), 154. https://doi.org/10.3390/educsci9020154

De Vos, W., & Pilot, A. (2001). Acids and Bases in Layers: The Stratal Structure of an Ancient Topic. *Journal of Chemical Education*, *78*(4), 494. https://doi.org/10.1021/ed078p494

Duit, R., Gropengießer, H., Kattmann, U., Komorek, M., & Parchmann, I. (2012). The Model of Educational Reconstruction: A Framework for Improving Teaching and Learning Science. In D. Jorde & J. Dillon (Eds.), *Cultural Perspectives in Science Education: Vol. 5. Science Education Research and Practice in Europe: Retrospective and Prospective* (pp. 13–37). SensePublishers.

Erduran, S., & Kaya, E. (2019). *Transforming Teacher Education Through the Epistemic Core of Chemistry: Empirical Evidence and Practical Strategies*. *Science*. Springer International Publishing AG. https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=5788445

Furió-Más, C., Luisa Calatayud, M., Guisasola, J., & Furió-Gómez, C. (2005). How are the Concepts and Theories of Acid–Base Reactions Presented? Chemistry in

Textbooks and as Presented by Teachers. *International Journal of Science Education*, *27*(11), 1337–1358. https://doi.org/10.1080/09500690500102896

Ghosh, A., & Berg, S. (2014). *Arrow pushing in inorganic chemistry: A logical approach to the chemistry of the main-group elements*. John Wiley & Sons Inc. https://doi.org/10.1002/9781118924525

Hand, B., & Treagust, D. F. (1988). Application of a conceptual conflict teaching strategy to enhance student learning of acids and bases. *Research in Science Education, 18*, 53–63

Hoe, K. Y., & Subramaniam, R. (2016). On the prevalence of alternative conceptions on acid–base chemistry among secondary students: insights from cognitive and confidence measures. *Chemistry Education Research and Practice*, *17*(2), 263–282. https://doi.org/10.1039/C5RP00146C

Jiménez-Liso, M. R., López-Banet, L., & Dillon, J. (2020). Changing How We Teach Acid-Base Chemistry: A Proposal Grounded in Studies of the History and Nature of Science Education. *Science & Education*, 1–25. https://doi.org/10.1007/s11191-020-00142-6

Justi, R., & Gilbert, J. (1999). A cause of ahistorical science teaching: Use of hybrid models. *Science Education*, *83*(2), 163–177. https://doi.org/10.1002/(SICI)1098-237X(199903)83:2<163::AID-SCE5>3.0.CO;2-I

Justi, R., & Gilbert, J. (2000). History and philosophy of science through models: some challenges in the case of 'the atom'. *International Journal of Science Education*, *22*(9), 993–1009. https://doi.org/10.1080/095006900416875

Kousathana, M., Demerouti, M., & Tsaparlis, G. (2005). Instructional Misconceptions in Acid-Base Equilibria: An Analysis from a History and Philosophy of Science Perspective. *Science & Education*, *14*(2), 173–193. https://doi.org/10.1007/s11191-005-5719-9

Krebs, R. E. (2023). Designing an Acid-Base Strength Interactive Screen Experiment (ISE): App Development with Upper Secondary School Students. In M. Rusek, M. Thóthová, & D. Koperová (Eds.), *Project-based and other student-activation strategies and issues in science education xx.* (pp. 2–11). Charles University Prague.

Krebs, R. E., & Hofer, E. (2022). Von den „scharfen Wässern" zu den „harten Säuren" – ein Streifzug durch die Geschichte der Säure-Base-Modelle. *Plus Lucis*(3), 8–11.

Krebs, R. E., Hofer, E., & Lembens, A. (2022). „Protons as the main drivers of a chemical reaction?": Educational Reconstruction of the Brønsted-Lowry Acid-Base Concept for Upper Secondary School. *CHEMKON*, *29*, 1–7. https://doi.org/10.1002/ckon.202200045

Krebs, R. E., & Lembens, A. (2021). Developing Key Ideas to Teach 'Acids' & 'Bases' in Upper Secondary Schools. In M. Rusek, M. Tóthová, & K. Vojíř (Chairs), *project-based education and other activating strategies in science education xviii,* Prague.

Krebs, R. E., & Lembens, A. (2024). "Acids are those dangerous green liquids, and

what's a base?'': Evaluating upper secondary students' 'acceptance' of a learner-appropriate approach to teaching about acid-base reactions. *Progress in Science Education*, *7*(1), 26–39. https://doi.org/10.25321/prise.2024.1423

Krebs, R. E., Rost, M., & Lembens, A. (2023). Developing and evaluating a multiple-choice knowledge test about Brønsted-Lowry acid-base reactions for upper secondary school students. *Chemistry Teacher International*, 5(2), 177-188. https://doi.org/10.1515/cti-2022-0038

Kuckartz, U. (2014). *Qualitative text analysis: A guide to methods, practice & using software* (A. Mcwhertor, Trans.). SAGE. http://gbv.eblib.com/patron/FullRecord.aspx?p=1633856

Lancaster, K., Malley, C., Gruneich, B., Loeblein, P., Moore, E. B., Parson, R., & Perkins, K. (2021). *Acid-base solutions*. https://phet.colorado.edu/sims/html/acid-base-solutions/latest/acid-base-solutions_en.html

Lembens, A., Hammerschmid, S., Jaklin-Farcher, S., Nosko, C., & Reiter, K. (2019). Textbooks as source for conceptional confusion in teaching and learning 'acids and bases' in lower secondary school. *Chemistry Teacher International*, *1*(2), 20180029. https://doi.org/10.1515/cti-2018-0029

Lewis, G. N. (1923). *Valence and the structure of atoms and molecules*. Chemical Catalog Comp.

Pan, H., & Henriques, L. (2015). Students' Alternate Conceptions on Acids and Bases. *School Science and Mathematics*, *115*(5), 237–243. https://doi.org/10.1111/ssm.12124

Sieve, B. F., & Bittorf, R. M. (2016). Protonenübergang oder Elektronenpaarübertragung? Säure-Base-Reaktionen sachgerecht darstellen. *Naturwissenschaften Im Unterricht / Chemie*, *155*, 47–48.

Stone, E. R. (2010). t-Test, Independent Samples. In N. J. Salkind (Ed.), *Encyclopedia of research design* (pp. 1560–1565). SAGE.

Taber, K. S. (2013). Revisiting the chemistry triplet: drawing upon the nature of chemical knowledge and the psychology of learning to inform chemistry education. *Chem. Educ. Res. Pract.*, *14*(2), 156–168.

Ültay, N., & Calik, M. (2016). A Comparison of Different Teaching Designs of 'Acids and Bases' Subject. *EURASIA Journal of Mathematics, Science & Technology Education*, *12*(1). https://doi.org/10.12973/eurasia.2016.1422a

Watson, S. W., Dubrovskiy, A. V., & Peters, M. L. (2020). Increasing chemistry students' knowledge, confidence, and conceptual understanding of pH using a collaborative computer pH simulation. *Chem. Educ. Res. Pract.*, *21*(2), 528–535. https://doi.org/10.1039/c9rp00235a

Wilson, M. (2004). *Constructing Measures*. Routledge. https://doi.org/10.4324/9781410611697

Wu, M., Tam, H. P., & Jen, T.-H. (2016). *Educational Measurement for Applied Researchers*. Springer Singapore. https://doi.org/10.1007/978-981-10-3302-5

Language And Middle-School Science: What Do Children And Teachers Gain By Implementing Semantic Waves In Teaching?

Suzanne Rønhøj Schjøtt, Dorte Maiken Lohse, Anna-Vera Meidell Sigsgaard

University College of Copenhagen, Denmark

This paper presents a small part of a three-year research project integrating language awareness in science lessons in Danish middle-school science classes. The project aims at making knowledge hidden in the language of science visible to all students, as many students, especially second-language learners, need more help understanding the abstract science language in a typical science class. In this part of the project, we followed a 6th-grade from a Danish primary school learning about the periodic table of elements and the molecules that build our material world. In collaboration with the teacher, we adjusted the lesson plan so that the teaching followed a progression plan from everyday language to a more appropriate abstract scientific language. In this paper, we apply semantic gravity analysis from Legitimation Code Theory (LCT) to student-student and teacher-student conversations and a written text. Based on semantic gravity analysis of various teaching situations, we argue that a student's understanding of science subjects depends on the teacher's ability to link common-sense knowledge with scientific knowledge and the teacher's ability to create coherence through special attention to the different strengths of semantic gravity, i.e. experiments and scientific explanations in conversations.

Keywords: Legitimation code theory (LCT).

Introduction

It is often assumed that students learn correct vocabulary through teachers' explanations, classroom discussions, different types of assignments and experiments, feedback from assessments and evaluation of their oral and written work. However, learning science is more complicated; you must also understand and use verbal discourse, and graphical and visual representations related to the subject (Lemke, 1998) and be able to explain an abstract scientific theory that can be understood as a generalization, rule, or concept that is not related to real-life actions or objects (Hyllested, 2020). In addition, students must also unlearn pre-learned, everyday understandings as they often differ from with the actual scientific explanations (Vosniadou, 2019). Thus, teaching science demands that

the teacher has not just scientific knowledge, but also knowledge about how we use language in science.

Research shows second-language learners struggle with science, often scoring lower grades (Beuchert et al., 2020). One of the reasons for the observed differences is that second-language learners are not only learning science and the language of science, which is new for all students; they are also learning it through a non-native language (Hammond & Gibbons, 2001). In traditional science teaching, many science teachers spend most of their attention on scientific or subject-specific words related to the topic (Jacobsen & Mulvad, 2022) and do not connect the students' everyday languages and understanding to the scientific explanations, making it difficult to understand the topics peculiarity. We argue that this blindness in science creates a difficult learning path, especially for second-language learners and lowers the academic level of all students. Thus, we argue that focusing on both the scientific language within the subject and students' pre-professional language would support most students (Derewianka & Jones, 2012). To meet this challenge, we have been exploring how teachers' awareness of so-called *semantic waves* (Maton, 2013) in middle school science classes, can affect their students' learning. The theoretical background is within 'Legitimate Code Theory' (Maton, 2014b), which provides analytical concepts for exploring how different kinds of knowledge are valued in various social contexts (Maton, 2014b). One such concept, *semantic gravity*, refers to how close meaning is tied to the context in which it occurs (Maton, 2014a; Meidell Sigsgaard & Jacobsen, 2021). Stronger semantic gravity (SG+) is when the understanding is closely connected to the actual context, while weaker semantic gravity (SG-) is associated with less connection to any specific context, associated instead with more general and abstract knowledge.

One of the problems semantic gravity analyses can reveal is that teachers can have challenges scaffolding "up", i.e. making it possible for students to transform knowledge tied to a specific context (e.g. when performing an experiment) to more general, scientific understandings of the phenomenon or principle (Georgiou, 2015). Thus, we hypothesize that introducing teachers' to the concept of semantic gravity, and 'semantic waving', i.e., moving back and forth between stronger and weaker context-specific meanings, is essential for creating a strong learning environment for most students, as demonstrated in other educational contexts (Blackie, 2014; Brooke, 2017; Clarence & McKenna, 2017; Kirk, 2017; Macnaught et al., 2013) and could lead to more equal learning outcomes for all students. Based on analyses of teacher-student conversations, this study examines what happens when the teacher's lessons are planned from a language-pedagogy perspective (Derewianka, 1990; Gibbons, 2003; Windschitl et al., 2018) with the aim of students develop appropriate language to communicate about the scientific topic. We analyze conversations from the perspective of

semantic gravity, especially as students are building skills towards a scientific understanding of the topic.

Methods

In this project, we follow a 6th-grade at a public school in Copenhagen, Denmark, for one school year (2022-2023). The class has 27 students aged 11-13, and one-third second-language learners. We collaborated with their science teacher, to plan three topics of study based on the teaching-learning cycle (Derewianka, 1990; Martin & Rose, 2012; Martin & Rothery, 1980). This approach allows for a progression in the language demands of the activities, developing both the student' scientific understandings and their awareness of the language related to the subject. We observed and video-recorded the class during each of the three science topics: the periodic table, magnetism, and human anatomy. The data included observation notes, video recordings of the classroom, teacher and student conversations, students' group work and experiments. In addition, we interviewed most students at the end of each topic and held collaborative meetings with the teacher after each observed lesson to discuss what happened. In this paper, we refer to data and findings gained during the first topic of study 'the periodic table of elements', in which we observed three 90-minute lessons out of a total of nine. We observed the first, second and fourth lesson and have a written text that the students made with their teacher in the third lesson.

In the following section, we apply a semantic gravity analysis to 1) the teaching plan from the first three lessons, 2) a few selected parts of the teacher-students and student-student conversations from the first and second observed lessons, and 3) the text the students and teacher jointly constructed together in lesson 3. These analyses will be supplemented by statements collected from the teacher and the students in their final interview. For our analysis of 1-3 we are inspired by previous work using semantic gravity in educational contexts, differentiating between three strengths of semantic gravity (Kirk, 2017; Meidell Sigsgaard & Jacobsen, 2020, 2021). The most context-dependent meanings correspond with strongest semantic gravity (SG+). Meanings that are not directly dependent on any specific context correspond with the weakest semantic gravity (SG-). And meanings that generalize based on several examples correspond with intermediate strength semantic gravity (SG0).

In the context of our study, stronger semantic gravity (SG+) often corresponds directly to the experiment the student did in lesson 1 or the activity with words and a model of the experimental design that the students and teacher worked with in lesson 2. It is closely related to what the students did, saw and experienced in the experiment. An intermediate strength semantic gravity (SG0) corresponds with something more general about a scientific topic. These often generalize or summarize students' experiences from their experiments but without being

a precise scientific explanation. Intermediate strength semantic gravity differs from the stronger semantic gravity in that the meaning being expressed comprises of more accurate knowledge than referring directly to the experiment or the exercise such as when making a generalization based on several observations or when suggesting a rule / hypothesis. Weaker semantic gravity (SG-) occurs in our data when the teacher or students talk about a scientific phenomenon or theory at a more abstract level, without directly linking to an experiment or exercise done in class. Conversations were chunked into turns (following conventions from conversation analysis, see e.g. Sacks et al., 1974), and turns into meaningful sections of text (i.e. clauses), which were then coded in terms of the three strengths of semantic gravity. Although a crucial part of teaching, turns/clauses not related to the scientific subject (e.g., the teacher's class management, asides, and other unrelated disturbances) are not coded with semantic gravity as these are considered off topic. Examples of text analyses of conversations can be seen in the following sections.

Data Analysis and Results

Section 1. Overall Semantic Gravity In The Teaching Plan Lesson 1-3

Through lesson 1-3 (see Appendix 1, Table 1) both the students and the teaching are moving from stronger semantic gravity (SG+) to weaker semantic gravity (SG-) i.e., from a high context-dependent meanings towards more abstract understandings of the experiment. In the specific topic, this relates to the students engaging with a beginning understanding of the atomic building blocks from the periodic table of elements. This was also the intention when planning the lessons, as we asked the students to start by observing an unexplained experiment, that we knew they would have little scientific knowledge about. This provided the students with a shared experience and something they could discuss later. However, in each lesson, we saw that the semantic gravity was at all three levels within conversations, but mainly with a high gravity in lesson 1, moving to a lover gravity in lesson 3, where the theory behind the experiment was described. Examples of conversations in lesson 1-3 will be shown and explained in detail in the following sections.

Section 2. Semantic Waves In Conversations From Lessons 1, 2 And 3

From the experimental phase in lesson 1, we observed a conversation with some of the better students that went like this as they reached the end of the experiment in lesson 1:

Student 1: Wow! That was really cool.

Student 2: The smoke is still down there.

Student 1: Yes, the smoke is still down there, it looks really cool.

Researcher: What do you think happened? It exploded. Do you want to try and say why?

Student 2: Well, I think it had a chemical reaction.

Student 1: Just like that. H_2O, that is air, right? No, what is it… What it is, that is made?

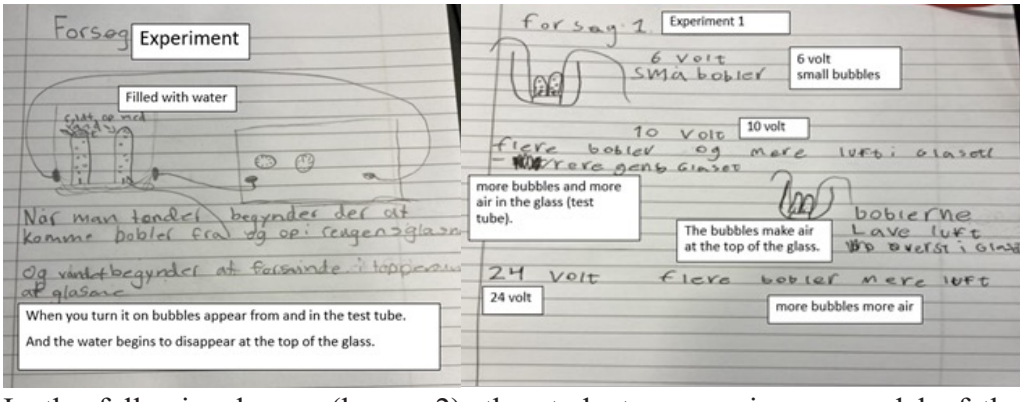
Researcher: There is a lot of things in the air. Oxygen. Oxygen among other things.

Student 1: I think, maybe, that the oxygen it made with the pressure and electricity has a chemical reaction with the fire.

Excerpt 1: Conversation between two students and a project researcher, after observing the cleaving of H_2O into constituent molecules (the electrolysis experiment). The text is coded as follows: strongest semantic gravity (SG+) in green, intermediate strength semantic gravity (SG0) in yellow, weakest semantic gravity (SG-) in blue. The original Danish text is available in Appendix 2B.

In excerpt 1 above, when asked what they thought happened in the experiment they had observed, the students start by answering with their observations (e.g. "the smoke is still down there"). As such their answer construes a meaning with stronger semantic gravity. When helped by the reseacher, their attempted explanation approaches a rule or hypothesis about the phenomenon: "I think, maybe, that the oxygen it made with the pressure and electricity has a chemical reaction with the fire". This understanding is less dependent on the immediate context and falls in the intermediate strength of semantic gravity (SG0), without yet expressing a more accurate scientific understanding of what happened. This makes sense, in that they were introduced to the theory only after completing the experiment. However, the conversation also shows that the students learn from the conversation and accumulate shared knowledge with each other and the teacher. Several examples of student's drawings and complimentary written explanations of the experiment can be seen in figure 1. Their texts show that their explanations are tightly linked to what they experienced and saw in the experiment.

Figure 1. Examples of students' drawings and explanations of the experiment in lesson 1.



In the following lesson (lesson 2), the students were given a model of the experimental setup, which they had observed in lesson one, as well as a list of words related to the experiment (technical words and process words). The students were asked to label the setup using the provided words and discuss their reasons for labelling as they did. See Appendix 1; Figure 1 for an example of the drawing and words students used to label the drawing with. The following excerpt is an example of a conversation during this activity:

Researcher:	So, a molecule is something that is joint together.
Student 1:	So different… What is it called? Elements? That could be joint together, joint together with different things.
Researcher:	Yes, so what is an element again?
Student 2:	Element, that is different substances.
Researcher:	Yes, and what is an element as well?
Student 2:	It can also be… different things.
Researcher:	Actually, it is a word we have here.
Student 2:	That? That word?
Student 1:	Is it… atom? Or..? Is it.. atom?
Researcher:	So, an element is also an atom…

Excerpt 2: Conversation between two students and a project researcher, while the students are labelling a drawn model of the electrolysis experiment (observed the day before). The text is coded with strongest semantic gravity (SG+) in green, intermediate strength semantic gravity (SG0) in yellow, weakest semantic gravity (SG-) in blue. The original Danish text is available in Appendix 2B.

Excerpt 2 shows how the students negotiate the meaning of the words and relate them to what happened in the experiment. During the exercise the students were

excited when discussing the words, their meanings and how they relate to the experiment they had observed in the first lesson. The excerpt shows that the students are moving between different levels of context dependent knowledge. The addition of provided words seems to encourage the students to use these words in their explanations of the experiment – and the understandings they are expressing, therefore, are more generalizing and scientifically appropriate (corresponding with an intermediate strength of semantic gravity, SG0), while still lacking full understanding of the phenomenon observed. Appendix 1; Figure 1 shows how students do link Oxygen to the air by labelling O_2 and air together. Although this is correct, they are not linking it as a gas in the test tube in the experiment as well, demonstrating only partial understanding.

Later in the same lesson, the teacher and the students went through the assignment together and drew the experimental setup on the board, while the teacher ensured all the students understood the words from the word list, by explaining each word and its placement on the drawing. This 18 min. long conversation included scientific explanations of words as well as the everyday meaning of them. E.g., for words like water, liquid, gas, oxygen, atoms, molecules (see excerpt 3 below).

Teacher:	And then we have a word called H_2O. Where did you place H_2O, Masha?
Student:	Uhm, H_2O is water. Is water.
Teacher:	Why don't we call it H_2O then? Why do we call it water? Why? I mean, if this is the same? Why do you think?
Student:	Water is a kind of the everyday way of saying it.
Teacher:	Yes.
Student:	And then H_2O is more like the technical way.
Teacher:	Well said! It is completely correct. It is the technical way. Could you say something else about this word H_2O? Fie?
Student:	H_2O is like hydrogen and oxygen mixed together.
Teacher:	Yes, that is correct. It is. Hydrogen and oxygen mixed together.

Excerpt 3: Conversation summing up the labelling activity between the teacher and one student at a time. The text is coded as before: strongest semantic gravity (SG+) in green, intermediate strength semantic gravity (SG0) in yellow, weakest semantic gravity (SG-) in blue. The original Danish text is available in Appendix 2B.

In excerpt 3 the students and the teacher are discussing what H_2O is. In this excerpt, students are able to express less context dependent understandings, "H_2O is water". However, they also demonstrate an ability to link this understanding with their more immediate context, i.e., when saying, "*Water* is a kind of the everyday way of saying it, and then H_2O is more like the technical way". The teacher acknowledges the students for their contributions, before following up with new questions, asking the students to think about everyday language versus the scientifically more correct terminology. In this way, the excerpt shows how through the conversation, everyday words, used to express context dependent meanings (SG+), are linked with reflections over these words and their different contexts (SG0) to the more scientifically correct explanations (SG-), i.e. when the students explain that "H_2O is like Hydrogen and Oxygen mixed together". These kinds of 'semantic waves' (Maton, 2014a) occurred often in this conversation, especially towards the end of the conversation (see Appendix 1; Figure 2). Although the students did not express an entirely correct explanation of the experiment, the excerpt shows them developing an awareness of the differences in everyday versus scientific terminology, as well as a better understanding the how matter (water) is made up of different elements, i.e., oxygen and hydrogen (H_2O).

The entire 18-minute-long conversation at the end of the lesson was analyzed from the perspective of semantic gravity (Appendix 1; Figure 2) and shows that the teacher-student conversation fluctuates between the three strengths of semantic gravity. In the beginning of the conversation, it is mainly related to the assignment (matching words and the experimental model), thus the knowledge in question has a stronger semantic gravity in that it relates directly to the context of the classroom. Later in the conversation, however, both the intermediate strength and the weakest strength of semantic gravity are activated by both the teacher and students' contributions to the conversation with more scientific explanations and references to atoms and the periodic table of elements, as excerpt 3 above shows.

Section 3. Semantic Waves In A Jointly Constructed Written Text About Electrolysis

In lesson three, students were asked to create a written text explaining the electrolysis experiment. The students made suggestions for each sentence, while the teacher mediated and helped to reformulate the students' suggestions to make the text more generically correct (see Appendix 2A). This cooperation around constructing at text is known as joint-construction (see e.g. Gibbons, 2009; Hunt, Anne, 1991; Martin & Rothery, 1980). The result is a dense, scientific text, only referencing the specific experiment observed occasionally (strongest semantic gravity as in, "it said puff when we lit the gas on fire"). Instead, as is appropriate for a scientific explanation, the text mostly makes meanings in the intermediate

strength of semantic gravity (generalizing meanings, such as "water does not conduct electric current very well") and weakest levels of semantic gravity, i.e. context-independent, scientifically accepted meanings such as "By applying electricity to the water, water is split into Hydrogen atoms and Oxygen atoms". Unfortunately, we did not observe this lesson, but according to the teacher in later interviews, the student-teacher conversation during the joint-construction always referred to the experiment, resulting in more semantic waving, than the produced text. As a result, the teacher commented that the students could follow the logic and produced a text beyond what they would be able to produce on their own. Having worked with the relevant words in the previous lesson, and having observed the experiment in the first lesson, was crucial for their ability to engage both with the language and the knowledge in a scientific context.

Conclusions

By planning the lessons in a chain of progressive language-use activities, with each exercise building up to the following exercise, it was possible for the students to see and develop appropriate language, as well as grow their scientific understanding. The student interviews conducted at the end of the topic (after the nine lessons) show the students demonstrating robust scientific insight into the topic of study (the periodic table of elements). They could explain what happened in their experiment, electrolysis of water. The interview showed that many students in good detail could explain how electricity could split water into oxygen (O_2) and hydrogen (H_2) gas by electrolysis. Most students, including many second language learners, showed a good understanding of the theme, and used scientifically appropriate language when discussing the experiment, molecules, atoms, and the periodic table of elements. In this way, they demonstrated abstract thinking by the end of the topic.

As stated in section 1-3 above, it was clear that the abstraction level was built slowly to allow all students to comprehend multiple levels within the language – both the everyday and scientific language. We started with an unexplained physics experiment to awaken the students' curiosity about the theme. Observing the experiment together gave students a common experience, to which they all could refer, and anchored their understandings of an otherwise very abstract way of understanding the world. This allowed the students to slowly move up the semantic wave, when working with the words and the experimental setup and later when writing the explanation in lessons 3.

Thus, the students did not get the explanation on what happened in the experiment before lesson 3, but they get to work on the meaning of each word in both an ordinary everyday understanding, and later with the scientific understanding of each word. Had we given the students the answers from the start, as often seen in ordinary 'cookbook' experiments, it is not sure that they would have gotten the

same out of the experiment and an in-depth understanding of the scientific and common words related to the experiment itself. Especially if the conversation did not move on all three levels of the semantic gravity continuum, and if the teacher did not help the students upwards towards less context-dependent meanings, which many students often appear to have difficulty with.

Furthermore, we argue that allowing for a slow progression in abstraction level is important to many students and not just for second language learners, as it allows all students to digest extra nuances within the spoken language, and at the same time, it helps to build scientific understanding and connection between experiments and theory if the teacher can 'wave' between different levels of context-dependent meanings. Thus, we suggest that analysis of teaching situations using semantic gravity could play an important role in developing teaching practices that ensure students' abilities to grasp abstract scientific content and connect these understandings with individual experiments and observations. Teachers' awareness of meanings having different levels of context-dependence, can help them plan activities and lessons in a way that scaffolds students' understandings and language development. When using semantic gravity as a planning tool, we must not only focus on weaker and stronger semantic gravity but also be aware of the waving itself and apply it in the teaching. This means that teachers must be able to make activities that connect stronger semantic gravity with weaker semantic gravity and, in doing so, help the students connect their common-sense understandings with the scientific language.

Based on the above results and discussion section, we argue that a student's understanding of science subjects depends on the teacher's ability to link common-sense knowledge with scientific knowledge. We suggest that reflecting on semantic gravity and discussing nuances in the spoken and scientific language can help many students in building a stronger scientific understanding of phenomena over time, and offers potential for deeper learning, where the students can connect the real world they experience with the scientific explanations they meet e.g., in textbooks. Semantic gravity provides a tool for teachers to help their students express themselves in a more scientifically correct way and gives them a tool for understanding which meanings students are understanding, and which ones they need help need help accessing. It is our hypothesis that giving teachers the perspective of semantic gravity can help more students develop better and more robust understandings of science.

Acknowledgement

This project has been generously supported by the Novo Nordisk Foundation with grant reference number NNF20OC0065719 in the call "Project grants for natural science education and informal learning environments – Fall 2020".

References

Beuchert, L., Christensen, V. T., & Jensen, S. V. (2020). *PISA Etnisk 2018: PISA 2018 med fokus på elever med indvandrerbaggrund* (p. 134). VIVE - Det Nationale Forsknings- og Analysecenter for Velfærd. www.vive.dk

Blackie, M. A. L. (2014). Creating semantic waves: Using Legitimation Code Theory as a tool to aid the teaching of chemistry. *Chemistry Education Research and Practice, 15*(4), 462–469. https://doi.org/10.1039/C4RP00147H

Brooke, M. (2017). Using 'Semantic Waves' to Guide Students Through the Research Process: From Adopting a Stance to Sound Cohesive Academic Writing. *Asian Journal of the Scholarship of Teaching and Learning, 7*(1), 37–66.

Clarence, S., & McKenna, S. (2017). Developing academic literacies through understanding the nature of disciplinary knowledge. *London Review of Education, 15*(1), 38–49. https://doi.org/10.18546/LRE.15.1.04

Derewianka, B. (1990). Rocks in the Head: Children and the Language of Geology. In R. Carter (Ed.), *Knowledge about language and the Curriculum* (pp. 197–215). Hodder & Stoughton.

Derewianka, B., & Jones, P. (2012). *Teaching Language in context*. Oxford University Press.

Georgiou, H. (2015). Putting physics knowledge in the hot seat: The semantics of student understandings of thermodynamics. In K. Maton, S. Hood, & S. Shay (Eds.), *Knowledge-building: Educational studies in Legitimation Code Theory* (pp. 176–192). Routledge. https://doi.org/10.4324/9781315672342-19

Gibbons, P. (2003). Mediating Language Learning: Teacher Interactions with ESL Students in a Content-Based Classroom. *TESOL Quarterly, 37*(2), 247. https://doi.org/10.2307/3588504

Gibbons, P. (2009*). English Learners, Academic Literacy, and Thinking: Learning in the Challenge Zone.* Heinemann.

Hammond, J., & Gibbons, P. (2001). What is scaffolding? In J. Hammond (Ed.), *Scaffolding: Teaching and learning in language and literacy education* (pp. 1–14). PETA.

Hunt, Anne. (1991). *Negotiation in joint construction: Teaching literacy in early childhood* [Honor's Thesis]. University of Sydney.

Hyllested, T. (2020). *Natur/teknologididaktik: Naturfagsundervisning for begyndere*. Hans Reitzel.

Jacobsen, G. K., & Mulvad, R. (2022). *Ind i naturfag—Fra tekster til naturfaglig kompetence*. Akademisk Forlag.

Kirk, S. (2017). Waves of Reflection: Seeing Knowledges in academic writing. In J. Kemp (Ed.), *EAP in a rapidly changing landscape: Issues, challenges and solutions* (pp. 109–118). Garnet Publishing. https://www.researchgate.net/publication/310166189_Waves_of_Reflection_seeing_knowledges_in_academic_writing_Draft

Lemke, J. L. (1998). Multiplying meaning: Visual and verbal semiotics in scientific

text. In J. R. Martin & R. Veel (Eds.), *Reading Science* (pp. 87–113). Routledge. https://www.researchgate.net/publication/246905867

Macnaught, L., Maton, K., Martin, J. R., & Matruglio, E. (2013). Jointly Constructing Semantic Waves: Implications for teacher training. *Linguistics and Education, 24*, 50–63. https://doi.org/10.1016/j.linged.2012.11.008

Martin, J. R., & Rose, D. (2012). *Learning to Write, Reading to Learn: Genre, Knowledge and Pedagogy in the Sydney School*. Equinox.

Martin, J. R., & Rothery, J. (1980). *Write it Right* (Working Papers in Linguistics*)*. Linguistics Department, University of Sydney.

Maton, K. (2013). Making semantic waves: A key to cumulative knowledge-building. *Linguistics and Education, 24*(1), 8–22. https://doi.org/10.1016/j.linged.2012.11.005

Maton, K. (2014a). Building powerful knowledge: The significance of semantic waves. In E. Rata & B. Barrett (Eds.), *Knowledge and the Future of the Curriculum: International studies in social realism* (Issue Young 2012, pp. 181–197). Palgrave Macmillan. https://doi.org/10.1057/9781137429261_12

Maton, K. (2014b). Knowledge and knowers: Towards a realist sociology of education. In Knowledge and Knowers: *Towards a realist sociology of education*. Routledge. https://doi.org/10.4324/9780203885734

Meidell Sigsgaard, A.-V., & Jacobsen, S. K. (2020). At bølge i en akademisk analyse. *Globe: A Journal of Language, Culture and Communication, 101*(10), 82–101.

Meidell Sigsgaard, A.-V., & Jacobsen, S. K. (2021). Semantiske bølger: Analytisk redskab og didaktisk princip. In S. K. Knudsen & L. Wulff (Eds.), *Kom ind i sproget: Flersprogede elever i fagundervisningen* (pp. 73–90). Akademisk Forlag.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language, 50*(4), 696–735. https://doi.org/10.2307/412243

Vosniadou, S. (2019). The Development of Students' Understanding of Science. *Frontiers in Education, 4*. https://doi.org/10.3389/feduc.2019.00032

Windschitl, M., Thompson, J. J., & Braaten, M. L. (2018). *Ambitious science teaching*. Harvard University Press.

Appendix 1

Table 1. The teaching plan for lesson 1-3, where we collaborated on the content.

Time frame	What happened in the lesson	What did the students do?
Lesson 1* 90 minutes	The students conducted an unexplained experiment, 'hydrolysis of water,' in groups. The teacher explained the setup and how to use the given materials. She did not go into detail about the theory behind the experiment.	The students conducted the experiment and had to write down and draw what happened in the experiment and try to explain what happened in their own words. Examples of student-drawn explanations are shown in Figure 1 in the text.
Lesson 2* 90 minutes	The students were given a model of the experimental setup and a list of relevant scientific and process words they had to place and explain in relation to the experiment they did in the previous lesson. Afterwards, the teacher drew the experimental setup model on the classroom board, followed by a teacher-student conversation explaining the scientific and everyday meaning of words and relevant theory regarding the experiment.	Firstly, the students had to argue about the meaning of each word they were given and place it correctly in the model. Secondly, they had to write and draw in their notebook as the teacher went through the setup and, in collaboration with the students, explained the words. E.g., describing words like solid, water and gas, combining everyday use of the word and the scientific understanding of each word. This work can be seen in Appendix 1, Figure 1.
Lesson 3 90 minutes	The students and teacher continued the work from last time as they hadn't finished it. Then, the students had time to write, in their own words, their explanation of what happened in the experiment, followed by a joint construction of a text explaining electrolysis, with the teacher guiding the process regarding what words and sentence structures to use.	Firstly, students work in groups to write their own text about the theory behind the experiment with the teacher's help. Secondly, they reviewed the theory and made a new text with the teacher and their classmates. This text is seen in Appendix 2A.

*Observed lessons. Lesson four that we also collaborated about and observed is not included in the above table or analysis. In the rest of the nine lessons the teacher planned the content.

Figure 1. Models from lesson 2. Two examples of words students in groups added to a model of the experiment.

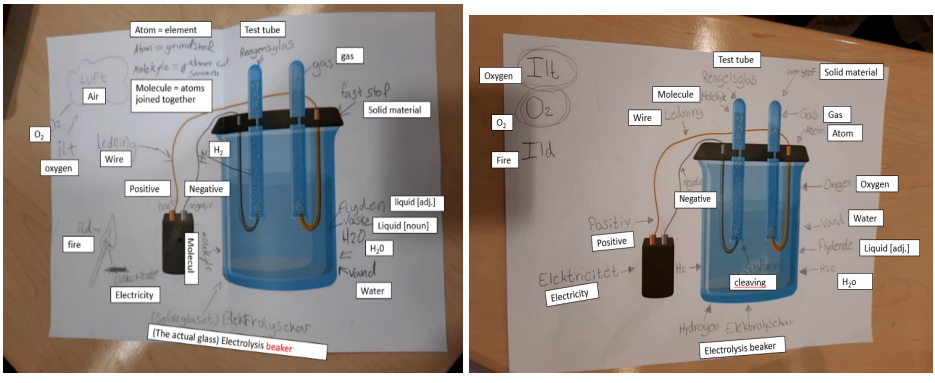
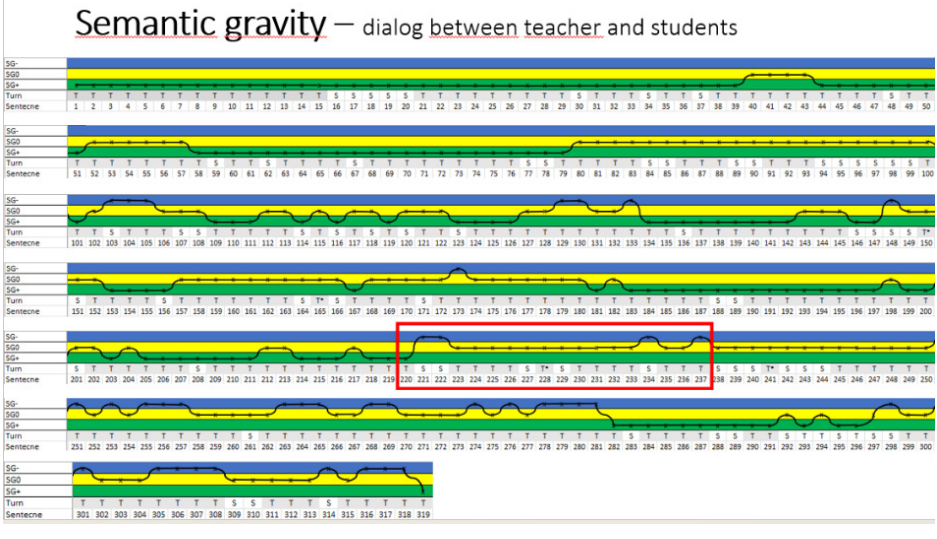


Figure 2. The figure shows the semantic gravity coding for the 18 mins of classroom conversation led by the teacher summarizing the labelling activity in lesson 2.



Each turn by a speaker (T for teacher, S for student) has been divided into sentences, which were then coded according to the three strengths of semantic gravity. Strongest semantic gravity in green, intermediate strength semantic gravity in yellow and weakest semantic gravity in blue. The red box denotes the turns demonstrated in excerpt 3.

Appendix 2

A: Semantic gravity analysis of the joint-constructed text from lesson 3. Our translation, followed by the Danish original**:** Weakest semantic gravity (SG-) is blue, intermediate strength semantic gravity (SG0) is yellow, and strongest semantic gravity (SG+) is green.

> 'With an electrolysis beaker, we can split water into atoms. Water is a molecule consisting of 2 Hydrogen and one Oxygen. We also call water H_2O. By applying electricity to the water, water is split into Hydrogen atoms and Oxygen atoms. Pure water is actually insulating. That is, water does not conduct electric current very well. Therefore, acid is added, which makes the electricity flow much better in the water. We also know Oxygen as oxygen. Since both Oxygen and Hydrogen are gases, the water H_2O, which is a liquid, becomes two gases. So both Hydrogen and Oxygen are a gas, but combined it becomes a liquid. We therefore produced two different gases in the two test tubes. In one glass there was Hydrogen gas and in the other Oxygen gas. It was much easier and faster to produce Hydrogen than Oxygen. This is because the water molecule consists of twice as many Hydrogen atoms as Oxygen atoms. H_2O is H-H-O. In addition to seeing twice as much gas in one test tube as in the other, we could also prove the two gases by setting them on fire. Since the Hydrogen atom is explosive, it said puff when we lit the gas on fire. While the flame flares up when we lit the Oxygen gas on fire. Fire lives off of Oxygen and the flames are strengthened by oxygen.'

> Original: 'Med et elektrolysekar kan vi spalte vand i atomer. Vand er et molekyle, der består af 2 Hydrogen og et Oxygen. Vi kalder også vand for H_2O. Ved at sætte strøm til vandet, spaltes vand til Hydrogen-atomer og Oxygen-atomer. Rent vand er i virkeligheden isolerende. Det vil sige, at vand ikke leder elektrisk strøm særligt godt. Derfor tilsættes syre, som gør at elektriciteten strømmer meget bedre i vandet. Vi kender også Oxygen som ilt. Da både Oxygen og Hydrogen er gasser, bliver vandet H_2O, der jo er en væske til to gasser. Så både Hydrogen og Oxygen er en gas, men sammensat bliver det til en væske. Vi fik derfor lavet to forskellige gasser i de to reagensglas. I det ene glas var der Hydrogen-gas og i det andet Oxygen-gas. Det var meget nemmere og hurtigere at producere Hydrogen end Oxygen. Det er fordi, at vand-molekylet består af dobbelt så mange Hydrogen-atomer end Oxygen-atomer. H_2O er H-H-O. Udover, at vi i reagensglasset kunne se dobbelt så meget gas i den ene end den anden, kunne vi også bevise de to gasser ved at sætte ild til dem. Da Hydrogen atomet er sprængfarligt, sagde det puf, da vi satte ild til gassen. Mens flammen bluser op, da vi satte ild til Oxygen-gassen. Ild lever af Oxygen og flammerne styrkes af ilt.'

B: Original Danish transcription of excerpts 1, 2 and 3, including semantic gravity coding: SG+ in green, SG0 in yellow and SG- in blue.

Excerpt 1:

Elev 1: Wow! Det var virkelig sejt.

Elev 2: Røgen er stadig dernede.

Elev 1: Ja, røgen er der stadig, det ser ret sejt ud.

Forsker: Hvad tror I, der skete der? Det eksploderede. Vil I prøve at sige hvorfor?

Elev 2: Jamen jeg tænker, at den havde en kemikal reaktion.

Elev 1: Bare lige sådan. H2O, det er luft, ikke? Nej. Hvad er det... Hvad er det, der er

lavet?

Forsker: Der er mange ting i luften. Oxygen. Oxygen blandt andet.

Elev 1: Jeg tror sådan, det oxygen, den har lavet med trykket og strøm, har en kemisk reaktion mod ilden

Excerpt 2:

Forsker: Så et molekyle, det er noget, der er sat sammen.

Elev1: Så forskellige… hvad er det, det hedder? Grundstoffer? Det kunne da godt være sat sammen, sat sammen af forskellige ting.

Forsker: Ja, så hvad er et grundstof igen.

Elev 2: Grundstof, det er forskellige stoffer.

Forsker: Ja, og hvad er et grundstof også?

Elev 2: Det kan også være... forskellige ting.

Forsker: Det er faktisk et ord, vi har her [peger på ord på papiret].

Elev 2: Den? Det ord?

Elev 1: Er det… atom? Eller…? Er det… atom?

Forsker: Så et grundstof, det er også et atom…

Excerpt 3:

Lærer: Og så har vi et ord, der hedder: H_2O. Hvor fik I placeret H_2O, Masha?

Elev: Øhm, H_2O er vand. Er vand.

Lærer: Hvorfor kalder vi det så ikke H_2O? Hvorfor kalder vi det vand? Hvorfor? Altså, hvis det der, det er samme? Hvorfor tror du?

Elev: Vand er lidt den daglige måde sige det på.

Lærer: Ja.

Elev: Og så H_2O er mere sådan den faglige.

Lærer: Ej, hvor det er godt sagt. Det er fuldstændig rigtigt. Det er den faglige måde. Kunne man sige noget andet om det her ord H_2O? Fie?

Elev: Altså H_2O det er ligesom sådan hydrogen og oxygen blandet sammen.

Lærer: Ja, det er rigtigt. Det er jo. H_2O er hydrogen og oxygen blandet sammen.

Writing Fiction to Expose Learning. An Adventure with Genetics and Evolution

José Manuel Ruvalcaba Cervantes, Irwing David Vásquez Cerqueda

Cinvestav, Monterrey, Nuevo León, Mexico

In science education, it is common to find examples that use narratives and storytelling to mediate science learning. Usually, the texts are read and analysed. Our work has another focus, science fiction writing to communicate learnings in science classes. We seek to answer the question what learning about genetic inheritance communicate high school students through writing science fiction short stories. To answer, we applied a workshop in which the students wrote science fiction short stories. The short story has to describe what appearance and the evolutionary changes humanity could experience. We suggested that writings use at least the keywords: phenotype, genotype, dominant and recessive allele. The workshop has three phases, and we handed out two orientation cards: 1) to recover previous learning and 2) to have at hand the structure of a short story. The results indicate that the writing of the short stories does not show what the students have learned. Instead of finding precise use of the keywords considered, students narrated catastrophic and dystopian scenarios in which humanity would have other characteristics. The results also reveal the potential to study alternative conceptions of evolution and probable relevance of the personal construction of conceptual orientations to mention keywords. We hope that this work provides empirical elements to discuss, in a broader project, the relationship between scientific imagination and scientific thinking skills.

Keywords: Science fiction, genetic, writing science.

Introduction

This work is part of a larger research project that discusses the relationship between scientific imagination and the development of scientific thinking skills. Here we seek to expose some findings of our project, specifically those that answer the question: what learning about genetic inheritance communicate high school students through writing science fiction short stories?

In science education, it is common to use narratives with stories of the key moments in the origin of scientific ideas, biographies, short stories and science fiction novels. The texts appear as didactic supports to be read and analysed to mediate learning. However, we take another approach, using writing and storytelling to communicate learning. We start from the assumption that

scientific imagination is a higher-order cognitive process socially mediated by symbolic tools and cultural signs (Ruvalcaba, et al., 2022). We hope short stories are a mediator that encourages imagining situations in which students explain keywords on which we focus our study.

Theoretical basis

Imagination as a socially mediated higher-order cognitive process means that imagining in scientific settings requires the support of symbolic mediators that regulate imaginative processes. That is, models, laws, and theories restrict the scientific imagination by making the imaginary situation - as in this case, the micro-stories of science fiction - a situation that is possibly real, possible unreal, impossibly real, impossible unreal (Gendler, 2010). It means that events, scenarios, or other imagined elements in science fiction stories are valued for their potential to happen in the physical or imaginary world as experienced by our senses. Thus, possible real indicates that the imagined situation has a probability of being observed or identified in the real world, while impossible real suggests that imagined has no place in reality. Possible unreal indicates that the imagined is likely to happen in the imagined conditions but not in physical reality, and impossible unreal implies that the events narrated in the imaginary scenario lack meaning both in imagined and actual conditions.

That scientific imagination is a socially mediated process implies that in science's teaching and learning processes, it is necessary to use means and tools to regulate what is imagined and communicated. The means and tools to regulate what is imagined and communicated allow students to become aware of what they know, make it explicit (as inner speech) and communicate it to others through external representation formats. The media and the tools guide students' imagination and thinking towards metacognition processes.

Writing in science classes requires the student to be aware of what they know and want to communicate; that is, to write, it is necessary to know what they are going to write about (Sanmartí, 2008). The nature of the construction of written discourses makes this representation format a favourable and relevant means to regulate imaginative processes in science learning. However, writing in science teaching is one of the most complex skills to develop (Sanmartí, 1999; 2008), which is why, in turn, it requires instruments that regulate the orientation of writing.

Writing science from a narrative approach reduces cognitive demand since narrative thinking is more natural, intuitive, and closer to daily reasoning than rational thinking (Gendler, 2010), which is typical of nominal scientific writing. Writing from a narrative perspective allows students to build semantic contexts to communicate ideas and arguments resulting from their learning. Familiarity with ideas and arguments favours that the student can abstract mental models

close to scientific ones at some point in the learning process.

Method

We carried out a science fiction creative writing workshop in which we asked the participants to write a short story that would answer the questions:

> What will happen to humanity in the centuries to come? That is, what appearance are we going to have? What transformations will we experience? How are we going to evolve?

We warn students that the short story written, even a science fiction story, should show learning about the keywords: genotype, phenotype, dominant and recessive allele.

We work with 22 Mexican high school students. The workshop lasted 180 minutes, and we organised it around the genetic inheritance content. The students had previously studied the topic in class with the purposes (DGB/DCA, 2018:20):

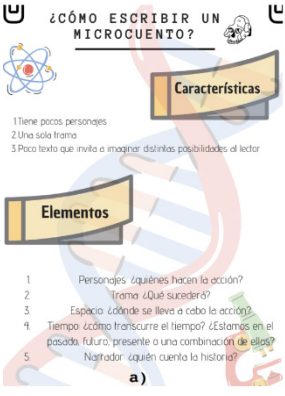
- "Explain the basic terms of heredity, favouring their creative development and identifying them in their environment […]
- Show mutation as a random process by critically reflecting on the impact on species."

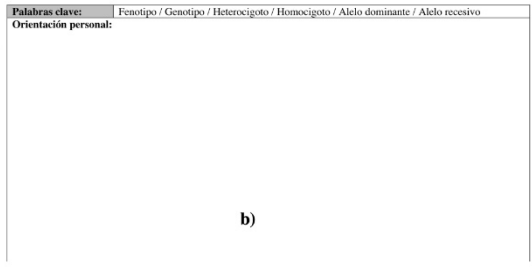
Instruments And Procedure

The writing workshop has three moments: 1) recover previous knowledge about keywords; we clarify students' doubts and explain it in plenary. 2) The students read a fragment of a science fiction story (Alvarado, 2016) to comment on the story's structure and how the author problematised humanity's transformation and evolution. 3) The students wrote (individually) and read (to the group) their short stories.

Two orientation cards (Figure 1) were provided to the students for the personal construction of orientation when writing the short stories. The first orientation card (Figure 1b) has a free and open format so each participant can build a conceptual guide with the keywords used in writing. The first card tally with the first moment of the workshop. Its design clarifies to the writer what is being written (Sanmartí, 2008). Each student constructed diagrams, notes, drawings, or diagrams with valuable ideas and personal meaning. The second card (Figure 1a) presents the structure and elements of a short story, highlighting the need for context, the macrostructure of the micro-story, and thinking about the audience. This card tally with the second moment of the workshop.

Figure 1 Orientation cards used in the workshop.





Results

We obtained 22 short stories of 1 page. Only 10 participants built a conceptual orientation sheet (Figure 1b) with the keywords recovered at the first moment of the workshop.

To account for how the learning of the keywords manifested, we first identified them in the students' writings. In Table 1, we appreciate the keywords used, separating the texts between students who built conceptual orientation and those who did not.

Table 1 Terms that appear in micro stories

Conceptually oriented		No conceptual orientation	
Keyword	Frequency	Keyword	Frequency
Recessive allele	-	Recessive allele	-
Dominant allele	1	Dominant allele	1
Phenotype	3	Phenotype	2
Genotype	5	Genotype	3

Some students did not mention all the keywords in the stories, and others used imprecisely, that is, with indefinite and incomplete exposition. We did not expect theoretical definitions, but we did expect a precise use of the terms following the narrated events due to the story's context. In Table 2, we find these evaluations.

Table 2 Uses of keywords

Conceptually oriented		No conceptual orientation	
Employment	Frequency	Employment	Frequency
Use 1 term	2	Use 1 term	1
Use 2 terms	3	Use 2 terms	3
Use 3 terms	-	Use 3 terms	-
Use 4 terms	-	Use 4 terms	-
Does not use key terms	5	Does not use key terms	8
Imprecise employment	5	Imprecise employment	4
Precise employment	-	Precise employment	-
Mention	5	Mention	4

Of the keywords mentioned, none is made accurately. Students make fantastic descriptions without linking key terms or others associated with genetic inheritance or evolution. Approximately 60% of students do not use keywords. They only narrate catastrophic and dystopian scenes that would force human survival, adaptation and extinction. Some participants list physiological characteristics ad-hoc to the imagined scenarios. In summary, it is not possible to determine what the students have learned.

Regarding the construction of personal conceptual orientation (POC), half of the students with POC mention keywords, and 33.33% of those who do not use POC mention keywords. The exact proportions are maintained when considering 1 and 2 terms. Of the students who built POC, 30% used two terms in their writing, compared to 25% of students who did not. In both cases, with POC and without POC, the genotype is the term with the highest number of mentions.

Writing science to communicate learning is more than enunciating vocabulary; developing a semantic context that provides meaning and theoretical structure to the words used is necessary. We assumed that writing science fiction stories would allow students to provide semantic context to communicate their learning; instead, we found a poor allusion to keywords of interest. However, we found elements that suggest other learnings linked to heredity, those about human evolution. The find is unexceptional because although the task requested of the students invites them to write the micro-story and asks them to use key and specific terms, the activity involves exposing ideas about human evolution.

The students expose biological changes that they imagine humans will evolve according to several conditions. The micro-stories narrate imaginary scenarios and stories to set the conditions of human evolution. In Table 3 are the micro-stories subjects that frame the conditions and human changes.

Table 3 Micro-stories subjects

Micro-story subject	Description
Environmental catastrophe	The students described scenarios where the Earth became uninhabitable. This could happen suddenly or gradually.
Cyberpunk	The plot concerns human evolution due to dystopian scenarios in which genetic engineering intervenes in the human genome. The result is decadent societies; humanity evolves until new species appear. After millennia, new species reveal themselves and create social chaos.
Catastrophic scenario	The plots narrate sudden catastrophic scenarios (due to environmental causes or wars) that modify the habitat of humanity. Consequently, one part of the species dies, and another adapts and survives.
Exodus	The narratives describe humanity's journey to colonize another planet, underground Earth, oceans, and space stations. People usually run away from Earth and its inhospitable conditions for humanity. The goal is to find habitats with conditions like the habitable Earth.
Genetics experiment	The plot revolves around genetic manipulation to create a hybrid of a human and a living being from another species or planet. The main objective of this hybrid is to adapt quickly to a new Earth's environmental conditions.
Human extinction	Plot the narrates the end of humankind.
Fantasy	Tales are freer and embrace the fantastic, with monsters and giants that are a branch of human evolution.
Alien fiction	These narratives give rise to contact between humans and beings from another planet. Hybrid species emerge in stories. Coexistence between 3 species ignites both cooperation and conflicts.

Discussion

Although students do not expose their learning on heredity and genetics and its application, a series of ideas and knowledge about human evolution appear. They list qualities and characteristics that humanity would develop according to the imagined scenarios. With this information, we reflect on the subjects: learning difficulties in genetic inheritance; teleology and determinism in evolution; didactic application contexts and difficulties in writing science.

Learning difficulties in genetic inheritance

Students wrote stories that lack mechanisms that explain the relationships between genes and human traits, as reported in the literature (Freidenreich et al., 2011; Puig et al., 2017). Most of the students' micro stories relate human evolution from an individual and not a population. The finding is explained because students conceive the gene from a deterministic and reduced position

and believe that the gene is the main driver of human evolution. For years, it has been reported (Flores Camacho et al., 2020) that students know what genes and what genetics studies are. However, they lack knowledge about the explanatory mechanisms of DNA replication, inheritance, and cell division.

In the same way, when writing about the characteristics of humans, according to the environmental conditions they are imagining, students express common errors (Stern et al., 2021; Puig et al., 2017): (1) the gene as the determining factor of the phenotype; (2) they believe that acquired characters are inherited.

In summary, it is not possible to determine what students have learned through writing science fiction stories, but some generalised alternative conceptions have been externalised.

Teleology and determinism in evolution

The stories involve sudden environmental catastrophes, social collapses, genetic experiments that go out of control, and scenarios that demand humanity an immediate adaptation to the new conditions. Such data indicate that students identify adaptation as a product of evolution and not adaptation because of natural selection.

In addition to the previous finding, students focusing on the characters and their characteristics implies manifesting determinism and teleology in human evolution and inheritance. The students express in their narratives that humans need to adapt to concrete and specific environmental conditions, so the evolution of humanity must be aimed at meeting the needs they refer to in the imagined scenarios. When populations manage to adapt, everything is determined by the gene. The findings are in line with what has been reported in the literature (Armenta, 2008; Flores Camacho et al., 2020; Freidenreich et al., 2011; González Gali, 2023; Puig et al., 2017; Stern et al., 2021): Students consider that the gene determines everything (from physiology to emotions), but it is possible to evolve according to the will of individuals to survive.

Didactic application contexts

The science fiction micro stories written by the students helped identify generalized conceptions: lack of mechanisms that explain the relationships between genes and human traits; conceiving evolution from an individual and not a population; conceiving the gene from a deterministic and reduced point; belief that the gene is the primary driver of human evolution; believe that the gene determining the phenotype; believe that acquired characters are inherited; identify adaptation as a product of evolution and not adaptation as a result of natural selection; deterministic and teleological ideas.

Different proposals (Stern et al., 2021) have emerged to address such ideas:

problem-solving, peer teaching, scientific inquiry, use of metaphors, visualization tools, manipulation of concrete material, working with bioinformatics tools, and many more. Regardless of the proposals, the results are similar (Flores Camacho et al., 2020; Stern et al., 2021): content knowledge is essential, but more is needed to understand genetics and overcome alternative conceptions, and students need help explaining the mechanisms linking genes and physical traits.

The relevance of didactic application contexts has been suggested for transferring learning from one situation to a new one (Puig et al., 2017), allowing students to apply what they have learned in solving new and different inheritance problems to consolidate learning. Other proposals seek that these didactic application contexts involve problems that require diverse representation formats (Flores Camacho et al., 2020) of genetics and its associated key terms.

It has been observed that a practical didactic application context to address alternative conceptions and problems to understand genetic inheritance is the study in the context of human genetics (Flores Camacho et al., 2020; Puig et al., 2017). Our work, although in a science fiction field, is in didactic contexts of application of human genetics; as we have mentioned, the context present here is not enough to account for what the students learned about genetic inheritance. However, it is still in line with what is reported in the literature since the contexts of didactic application with human genetics, even though they are more effective than other didactic contexts of application, are not enough to know if students understand the genetics models and successfully transfer them to different situations (Puig et al., 2017). In this sense, it is necessary to understand what happens in the context of applying genetics and the skills required to apply knowledge in the didactic context. In our case, it is necessary to direct research towards science fiction writing to understand the difficulties of applying what has been learned.

Difficulties in writing science

We found elements that suggest approving what Sanmartí (2008) said: writing science implies having an idea about what is written. Personal construction of conceptual orientations could support students to keep in mind the key ideas they write in communication processes. This is open the probability that POCs contribute to the conscious and deliberate acts of writing and communicating learning. Written skills are one of the most complex skills to develop in science education; POCs open up the possibility of supporting self-regulation and awareness in this challenging process.

The micro-stories narrate imaginary scenarios and stories to set the conditions of human evolution. Through micro-stories, the students deliver explanations of human evolution. These tales set out imaginary scenarios and narratives, giving different perspectives on the conditions that could shape our species. As

seen in Table 3, eight subjects reflect a cultural influence on how future changes in humanity are massively and culturally conceived under the idea of human evolution. We framed these subjects in risk narratives (Bilandzic et al., 2019). Therefore, the stories focus on the characters and their experientiality (Fludernik, 2010) due to drastic changes in the natural environment and the effects of these catastrophic changes over the characters.

Conclusion

The communication of learning the keywords is not appreciated in the short stories written by the students. We did not find elements that allow us to analyse learning, structures and causal relationships with which students endow scientific terms. In any case, using some terms at a rote level can serve as a basis for delving into skills that demonstrate a greater understanding of inheritance and expose alternative conceptions. Why learnings don't show up can be explained in several ways, from difficulties with creative writing to thinking that iconic and fictional content could replace poor understanding. However, we recognize the difficulty of writing a narrative story based on scientific content: The adaptation of scientific language to a story that its interpretation is through characters, actions, plot, and motives (Avraamidou & Osborne, 2009). It is difficult to explain because our work is limited: the workshop was not linked to the teaching and learning process, and fiction writing was not part of the process.

Our findings allow us to glimpse the relevance of some students' personal construction of conceptual orientations; such guidelines are self-regulatory prompts to at least mention keywords. This finding is relevant since scientific writing at school is an activity to self-regulate ideas and scientific thinking.

Acknowledgement

The first author thanks to CONAHCYT for the support with the Jóvenes Investigadoras e Investigadores por México program.

References

Armenta, M. C. (2008). Algunas ideas del alumnado de secundaria sobre conceptos básicos de genética. Enseñanza de las ciencias: revista de investigación y experiencias didácticas, 26(2). 227-244. DOI: 10.5565/rev/ensciencias.3677

Avraamidou, L., & Osborne, J. (2009). The role of narrative in communicating science. *International Journal of Science Education*, *31*(12), 1683-1707.

Alvarado Díaz. J.R. (2016). Nova Terra 5.0. en Las cuatro esquinas del universo. Rosetta. UNAM Micromegas.

Bilandzic, H., Kinnebrock, S., & Klingler, M. (2020). The emotional effects of science narratives: a theoretical framework. *Media and Communication*, *8*(1), 151-163.

Flores-Camacho, F., García-Rivera, B., Báez-Islas, A., Gallegos-Cázares, L., Calderón-Canales, E. (2020). Logros en la comprensión de temas de genética utilizando representaciones externas. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, *17*(3), 3101. doi: 10.25267/ Rev_Eureka_ensen_divulg_cienc.2020v17.i3.3101

Fludernik, M. (2010). Towards a 'natural' narratology (2nd. ed.). Routledge.

Freidenreich, H. B., Duncan, R. G., & Shea, N. (2011). Exploring Middle School Students' Understanding of Three Conceptual Models in Genetics. *International Journal of Science Education*, *33*(17), 2323-2349. https://doi.org/10.1080/09500693.2010.536997

Gendler, T. S. (2010). Intuition, imagination & philosophical methodology. Oxford University Press.

Gonzalez Galli, L. M. (2023). Analogías y enseñanza de la genética y la biología evolucionista. *Enseñanza de las Ciencias*, *41*(1), 63-78. https://doi.org/10.5565/rev/ensciencias.5615

Puig, B., Ageitos, N. & Jiménez-Aleixandre, M.P. (2017). Learning Gene Expression Through Modelling and Argumentation. *Sci & Educ*, 26, 1193–1222. https://doi.org/10.1007/s11191-017-9943-x

Sanmartí, N. (2008). Escribir para aprender ciencias. *Revista Aula de Innovación Educativa*, *175*, 29-32.

Sanmartí, N., Aymerich, M. I., & García, P. (1999). Hablar y escribir: Una condición necesaria para aprender ciencias. *Cuadernos de pedagogía*, (281), 54-58.

Stern, F., Kampourakis, K., Delaval, M., Müller, A. (2021). How are High-School Students' Teleological and Essentialist Conceptions Expressed in the Context of Genetics and What Can Teachers Do to Address Them? In: Haskel-Ittah, M., Yarden, A. (Eds), Genetics Education. Contributions from Biology Education Research. Springer, Cham. https://doi.org/10.1007/978-3-030-86051-6_9

Ruvalcaba Cervantes, J.M., Ruiz González, H.L., & Quintero Zazueta, R. (2022). Imaginación científica escolar: actividad social y cognitiva. *Sincronía*, 26. 81. 487-500. Doi: 10.32870/sincronia.axxvi.n81.24a22 http://sincronia.cucsh.udg.mx/pdf/81/487_500_2022a.pdf

Designing an Educational Escape Room: Evidence-Based Guidelines and Practical Advice

Georgios Villias

University of Cambridge, Cambridge, United Kingdom

Educational escape rooms (EERs) have been gaining momentum within the educational community since their first appearance a few years ago. Appearing in various forms (physical escape rooms, escape boxes, digital escape rooms), they managed to make it through the COVID-19 pandemic. Several different frameworks have been proposed for facilitating practitioners to create their own EER activities. This workshop, intents to familiarise its participants to the conceptual and design framework of an educational escape room as well as to present them examples of EER activities and different types of puzzles that have been thoroughly studied. It will also offer them the opportunity to actively engage in a biology-related EER activity, and brainstorm on the development of their very own educational escape room.

Keywords: Educational escape rooms, science education, design framework

Introduction and Theoretical Background

Educational Escape Rooms (EERs)

Educational escape rooms (EERs) are game-based activities that adopted the initial concept from the escape room industry and adapted it appropriately for use in an educational context. By offering immersive experiences that promote students' active participation, EERs have emerged as promising alternative approaches to fostering students' conceptual and skill-based learning (Nicholson, 2018). Being increasingly popular, the EERs' structure and usage has been systematically reviewed (Fotaris & Mastoras, 2019; Lathwesen & Belova, 2021; Taraldsen, Haara, Lysne, Jensen, & Jenssen, 2022; Veldkamp, van de Grint, Knippels, & van Joolingen, 2020). Their acceptance from students and teachers alike, supports their usability as learning environments for science education (Veldkamp, Knippels, & van Joolingen, 2021). As Lathwesen and Belova (2021) demonstrated while analysing the existing literature, the STEM domains in which most EER activities have been developed so far can be found primarily, but not exclusively, in chemistry education, mathematics and physics. The greatest interest in EER activities that could facilitate learning for university level students exists in the area of healthcare (pharmacy, nursing, medicine)

and computer science education. Taraldsen et al (2022) identified four fields of attention that EER activities usually focus on: exposure of participants to real-life scenarios, application of courses' curriculum content, development of learners' 21st century skills, and increase of motivation and engagement in the learning process.

Design frameworks

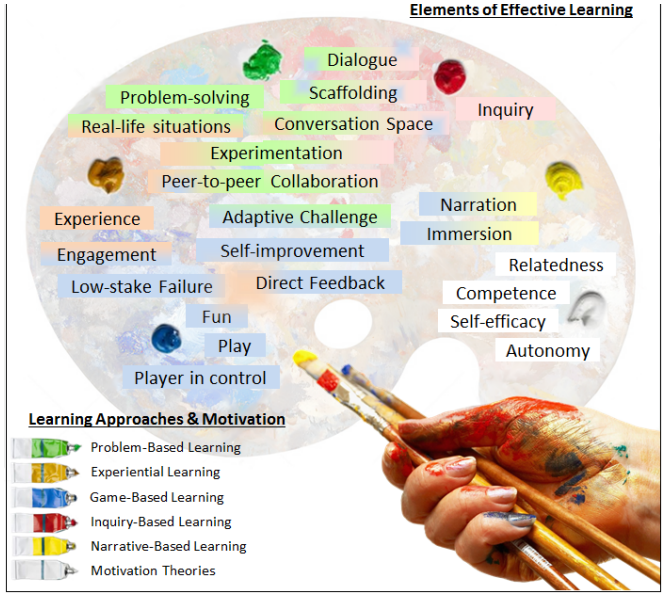
The educational community's demand for an easy to follow, comprehensive, and practical guide that will enable educators to take advantage of the learning benefits that EERs have to offer and create their own activities is raising. That is reflected on the existing literature about EERs. Several design frameworks and guidelines for the creation of educational escape rooms have been published from time to time (Clarke et al., 2017; Fotaris & Mastoras, 2022; Nicholson, 2018; Nicholson & Cable, 2021). Some of the EERs' design elements, like immersion, collaboration, and debriefing, have also been explored extensively (Veldkamp, Rebecca Niese, Heuvelmans, Knippels, & van Joolingen, 2022). Nevertheless, evidence-based guidelines and practical advice from experienced educators that have repeatedly designed and implemented EER activities is still priceless, especially for the early adopters.

Conceptual framework of an EER

Practitioners' and researchers' endeavours to tailor students' learning needs and maximise their learning opportunities, has led to the development of several teaching and learning approaches (e.g., problem-based, inquiry-based, experiential, game-based, narrative-based, dialogic/dialectic, etc.) that have been implemented in educational settings during the past few decades. Considering that in practice, dependent on learners' special characteristics, every learning approach may only benefit a part of the population, adopting a multi-dimensional approach to teaching and learning seems to be most suitable in producing better learning outcomes and reaching a larger proportion of the whole learners' population. A deeper look inside the design principles and the conceptual framework of an EER reveals their connection to several well-established educational methodologies (e.g., problem-based, inquiry-based, experiential, game-based, narrative-based, etc.), as well as motivational theories (e.g., self-determination theory, self-efficacy theory, etc.), justifying their learning potential (see Figure 1).

Figure 1. A. Conceptual framework of an EER activity. B, C. EER activities' "design"

A



B



C



In this painter's palette (A) are concentrated various elements of effectiveness that can be used in the learning process. Every colour tube represents a different learning approach. Education escape room activities have connections with all of the above learning approaches and motivation theories. Combinations of elements from all of them can be used dependent on the EERs' particular design and learning objectives. Considering the developer's experience and creativity, just like a painter's work, the result can be a simple drawing (B), clear, straightforward and easy to understand, or a unique masterpiece (C), complex, thrilling and imaginative.

Design guidelines

According to one of the latest and most complete design frameworks proposed by Fotaris and Mastoras (2022), there are several elements and parameters that need to be considered when designing an EER activity (e.g., demographic information about the participants' background, skill level, needs, and motivation; the activity's goal, learning objectives, constraints, required knowledge, group size, game type, playtime length, curriculum position, theme, setting, narrative, characters, puzzle types, puzzle designs, puzzle path, game flow, game assets, room layout, hint system, scoring system, introduction, rules, and reflection). Exploring thoroughly some of the features mentioned above, reflecting on the insights I gained from the findings of my personal research study, and building upon design guidelines provided by previous studies (Clarke et al., 2017; Fotaris & Mastoras, 2019, 2022; Nicholson, 2018; Nicholson & Cable, 2021; Veldkamp, Daemen, et al., 2020; Veldkamp et al., 2021, 2022; Veldkamp, van de Grint, et al., 2020), I recommend the following set of practical guidelines for designing EERs that focus both on the development of participants' 4C skills and content learning (Villias & Winterbottom, 2024):

1. forming teams of four or five members
2. using escape boxes
3. including physical objects as puzzle components
4. adopting a combination of linear and multi-linear puzzle pathways
5. designing self-guided puzzles that align well with the EERs' learning
6. adding layers to puzzles
7. increasing the number and the complexity of puzzle components
8. limiting the provision of scaffolding and hints to a minimum
9. incorporating meta-puzzles
10. challenging learners' pre-existing knowledge

Workshop's Activities

During the last 7 years, I have designed 8 educational escape rooms (5 for secondary and 3 for primary school students) and I have successfully implemented them in-person more than 50 times. I have also conducted research related to the design, implementation, and evaluation of EER activities' learning impact. Feeling confident that I have the experience to disseminate the knowledge I have acquired and desiring to go beyond a simple theoretical approach, I provided a series of hands-on activities.

1) After a brief welcoming at the workshop and an introductory presentation for educational escape rooms, the workshop's participants formed teams of 4-5 persons and actively engaged in an EER activity. The activity focused on a specific part of a STEM course's syllabus and had well-defined learning objectives. Experiencing first-hand what an educational escape room is all about, was an ideal initiation of the workshop's participants into this versatile educational tool.

2) An analysis of the presented EER activity followed in order to facilitate participants identify:

a) its overall structure (see Figure 2),
b) examples of puzzles (see Figure 3),
c) some general design guidelines (e.g., the use of an exciting, adventurous, or student age-related theme and narrative; the creation of a fun, playful, low-stake failure, and safe learning environment; the adoption of an inter-disciplinary approach to learning; the support of a culture of teamwork; the combination of tasks, challenges, and puzzles of different difficulty; the avoidance of puzzles with straightforward solutions), and
d) its special design elements (e.g., introduction to the story/narration, discovery of puzzles' resources and game material, number of puzzles and game steps, puzzles' connecting pathway and meta-puzzles, puzzle types, level of instructional guidance, facilitator's level of intervention, debriefing, content knowledge).

3) A design framework with specific steps was presented to participants (see Figure 4). Using an "instruction manual" that guided them through a step-by-step process, participants' working groups were encouraged to use a brainstorming approach, trying to conceptualise and outline their own syllabus-focused educational escape rooms.

4) As a final step, participants shared and discussed their ideas, inspiring others, but also getting valuable feedback from their peers.

Figure 2. An EER's special design elements and puzzles' overview

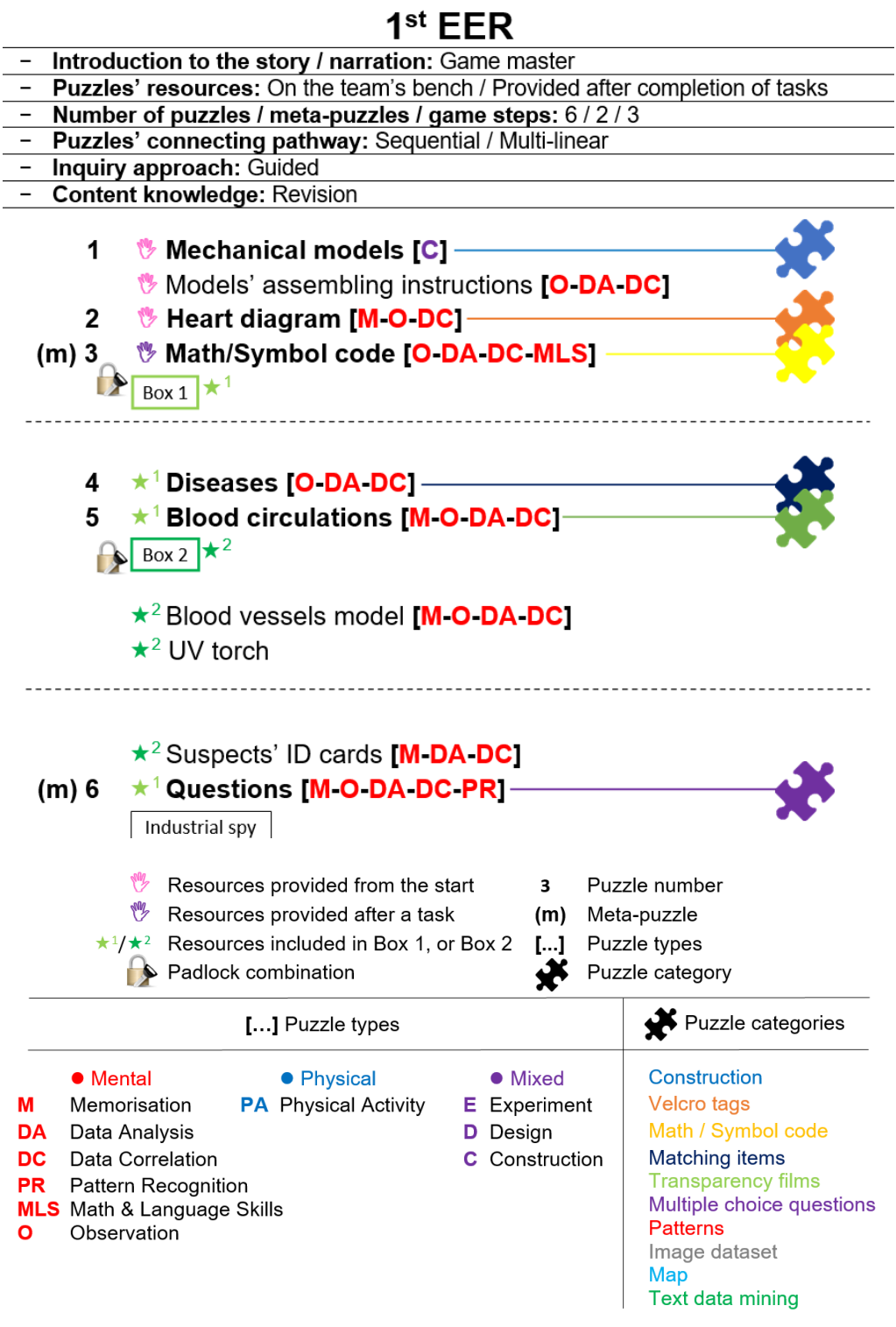
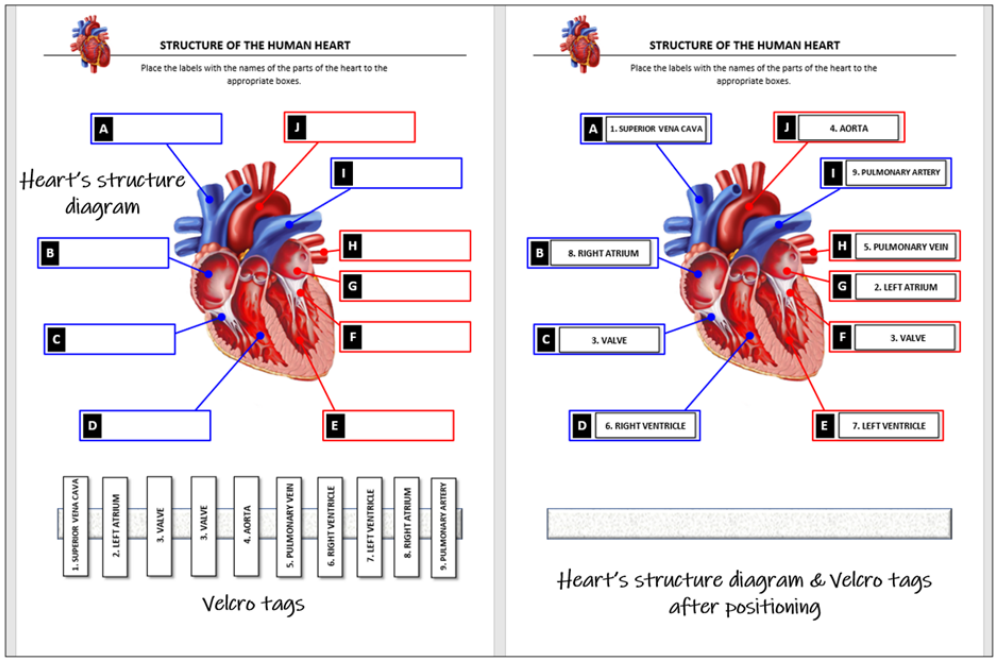
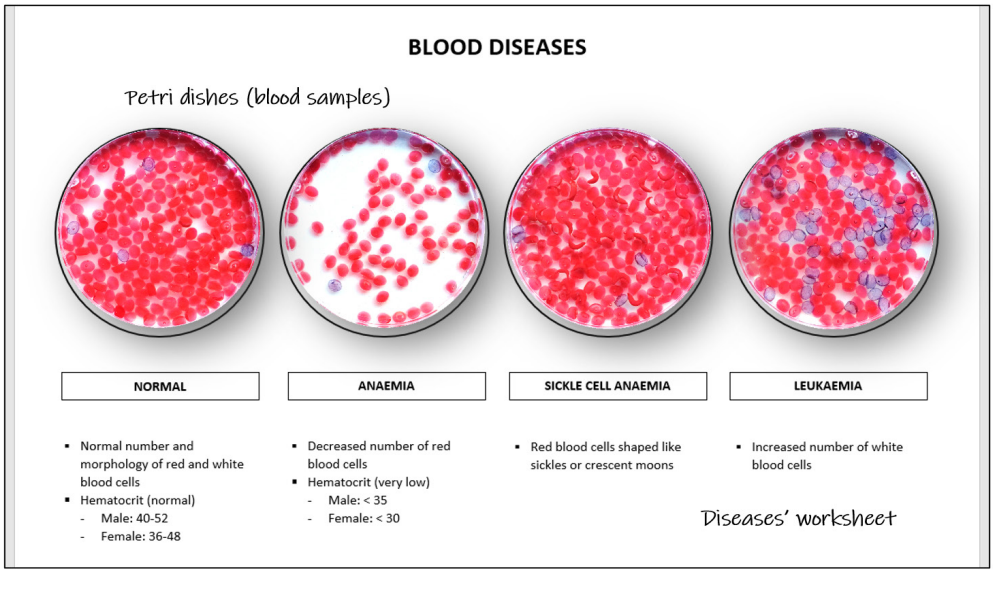


Figure 3. Resources of puzzle 2, 4, and 5.





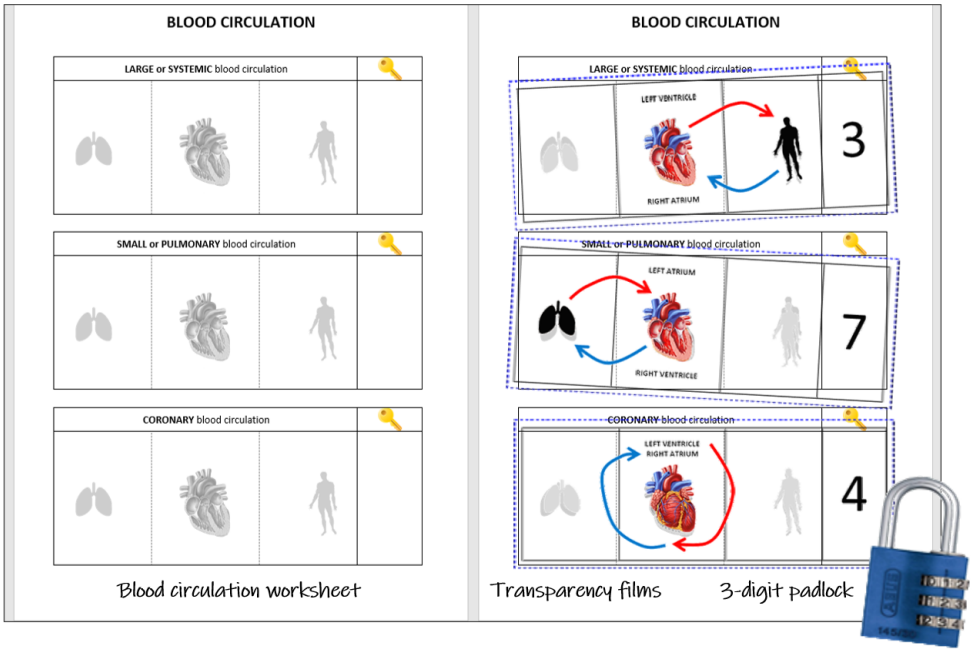
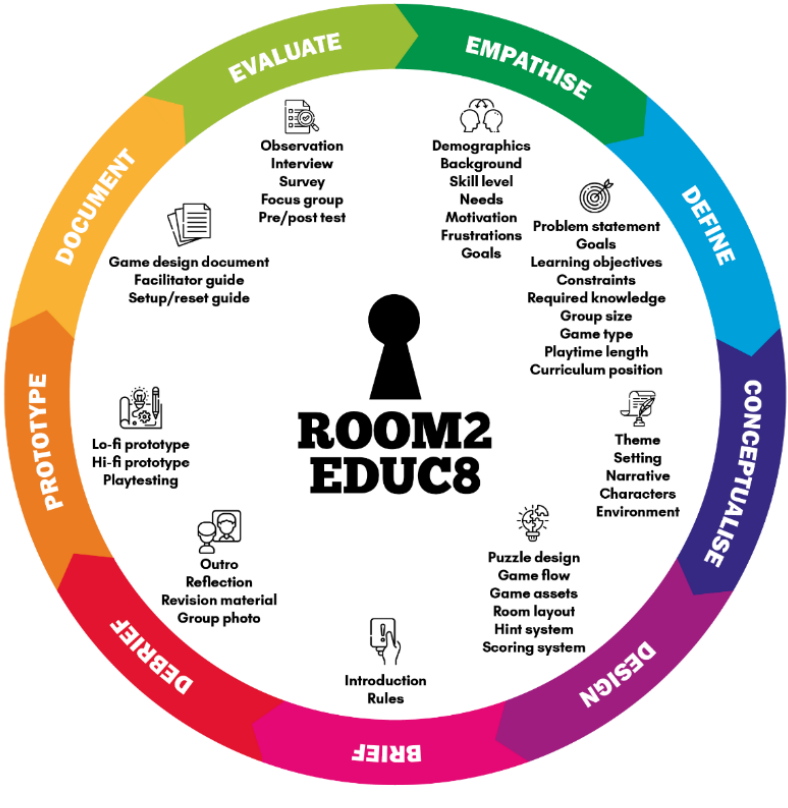


Figure 4. Room2Educ8 framework for designing EER activities (Fotaris & Mastoras, 2022)



References

Clarke, S., Peel, D. J., Arnab, S., Morini, L., Keegan, H., & Wood, O. (2017). escapED: A Framework for Creating Educational Escape Rooms and Interactive Games For Higher/Further Education. International Journal of Serious Games, 4(3), 73–86. https://doi.org/10.17083/ijsg.v4i3.180

Fotaris, P., & Mastoras, T. (2019). Escape rooms for learning: A systematic review. Proceedings of the European Conference on Games-Based Learning, 2019-October, 235–243. https://doi.org/10.34190/GBL.19.179

Fotaris, P., & Mastoras, T. (2022). Room2Educ8: A Framework for Creating Educational Escape Rooms Based on Design Thinking Principles. Education Sciences, 12(11), 768. https://doi.org/10.3390/educsci12110768

Lathwesen, C., & Belova, N. (2021). Escape rooms in stem teaching and learning—prospective field or declining trend? A literature review. Education Sciences, 11(6). https://doi.org/10.3390/educsci11060308

Nicholson, S., & Cable, L. (2021). Unlocking the potential of puzzle-based learning: designing escape rooms + games for the classroom. Corwin.

Nicholson, S. (2018). Creating Engaging Escape Rooms for the Classroom. Childhood Education, 94(1), 44–49. https://doi.org/10.1080/00094056.2018.1420363

Taraldsen, L. H., Haara, F. O., Lysne, M. S., Jensen, P. R., & Jenssen, E. S. (2022). A review on use of escape rooms in education–touching the void. Education Inquiry, 13(2), 169–184. https://doi.org/10.1080/20004508.2020.1860284

Veldkamp, A., Knippels, M. C. P. J., & van Joolingen, W. R. (2021). Beyond the Early Adopters: Escape Rooms in Science Education. Frontiers in Education, 6(April). https://doi.org/10.3389/feduc.2021.622860

Veldkamp, A., Rebecca Niese, J., Heuvelmans, M., Knippels, M. C. P. J., & van Joolingen, W. R. (2022). You escaped! How did you learn during gameplay? British Journal of Educational Technology, (July 2021), 1–29. https://doi.org/10.1111/bjet.13194

Veldkamp, A., van de Grint, L., Knippels, M. C. P. J., & van Joolingen, W. R. (2020). Escape education: A systematic review on escape rooms in education. Educational Research Review, 31(January), 100364. https://doi.org/10.1016/j.edurev.2020.100364

Villias, G., Winterbottom, M. (2024). Facilitating the Practice of 4C Skills in Biology Education Through Educational Escape Rooms. In: Korfiatis, K., Grace, M., Hammann, M. (eds) *Shaping the Future of Biological Education Research. Contributions from Biology Education Research. Springer, Cham*. https://doi.org/10.1007/978-3-031-44792-1_5



Strand 19

Teaching and Learning Science at the University Level

Federico Corni & May Lee

Strand Chairs/Co-editors

Foreword

The university years represent a critical period in the academic and professional development of students, especially those pursuing science-related disciplines. At this stage, the focus shifts from foundational knowledge to advanced concepts, research methodologies, and the application of scientific principles to real-world problems. The teaching and learning of science at the university level is, therefore, pivotal in preparing students for careers in science, technology, engineering, and mathematics (STEM) fields, as well as for roles in education, policy, and beyond.

Strand 19 of the ESERA 2023 proceedings, titled "Teaching and Learning Science at the University Level," brings together a collection of four proceedings that explore the complexities and opportunities of science education in higher education. These contributions offer valuable insights into the challenges faced by educators and students alike, as well as the innovative approaches being developed to enhance learning outcomes in university science courses.

The proceedings in this chapter address a range of topics, including curriculum design, instructional strategies, assessment practices, and the integration of research into teaching. They reflect a commitment to improving the quality of science education at the university level, ensuring that students are not only well-versed in scientific content but also equipped with the critical thinking, problem-solving, and communication skills necessary for success in their future endeavors.

Through these proceedings, readers are invited to engage with current research and discussions that are shaping the future of science education in universities. The insights provided here underscore the importance of fostering a learning environment that is both rigorous and supportive, encouraging students to explore, question, and innovate.

As you read through this chapter, we hope you find inspiration and practical ideas that can be applied to your own teaching or research practices. The knowledge shared in these pages is a testament to the dedication of educators and researchers who are committed to advancing science education at the highest levels.

We believe that the contributions in this chapter will not only inform but also inspire further exploration and dialogue on how to best support the learning journeys of university science students.

Culturally Responsive Methods in STEM Education

Nicoleta Gaciu[1], Edeltraud Gehrig[2]

[1]Oxford Brookes University, United Kingdom

[2]RheinMain University of Applied Sciences, Germany

STEM education is impacted not only by the cognitive overload of the content but also by the culture, the demands of the social and economic transformations and the communication of information in connection to context-specific inquiry. The integration of cultural heritage alongside the propositional and visual information in Culturally Responsive Methods (CRM) helps to reduce the obstacles that come up when learners have low prior knowledge, misconceptions, or insufficient interest in mathematics. The purpose of this article is to present the results from the first stage of a designed-based research project that is aimed at developing culturally responsive methods in mathematics undergraduate education that dynamically adapt to students' cultural heritage to facilitate the context-specific inquiry processes and to accommodate the culturally influenced and diverse ways students learn. The results confirm that CRM with direct adaption to individual cultural heritage, backgrounds and experience helped our students to achieve a concrete improvement in learning outcomes and creativity. Moreover, the students' cultural heritage embedded in CRM, is a strong factor that helps them to apply their knowledge to completely new situations. In addition, the results of the self-assessment suggest that learning is increased through the integration of self-chosen projects based on extended cultural heritage.

Keywords: culturally responsive methods, cultural heritage, project-based work

Introduction

The dialectic relationship between science education and culture refers to the interdependence between several factors, such as infrastructure, artefacts, cultural and scientific concepts, learning, creativity, and motivation, to name only a few. These factors mutually express each other, primarily because 'our interaction with nature is mediated through culture, our competencies for dealing with nature are cultural competencies for utilizing cultural means (concepts, artefacts) for dealing with nature.' (Ratner, 2008, p.5). Learning in science is an integrative process through which learners compare what is known to new knowledge and experiences, making sense of the multicultural world they live in and it is an 'essentially a social function devoted to dealing with cultural

information, and culturally mediated information about nature' (Ratner, 2008, p.5).

Integrating a student's culture into the curriculum has often been confined to activities that celebrate specific events designed to understand the culture of minority groups and not fully assisting and helping to meet the learning needs of multiculturally diverse groups of students. Since the development of Culturally Responsive Pedagogies (CRP) (Ladson-Billings, 1995) and Culturally Responsive Teaching (CRT) (Gay, 2000), the main teaching focus has primarily been on reducing the achievement gap between ethnic and racial minority groups of students and the dominant groups (Tate, 1995; Averill at al., 2009). Moreover, the CRP and CRT, as pedagogical models, have taken into consideration only the instructional aspects of acquiring knowledge in elementary education and humanities (Cunningham, 2001) and have lagged in science education (Dodo Seriki, 2018). However, in recent years, these models and methods have seen a change in acquiring diverse categories of knowledge in multicultural classrooms, focusing on how to teach all students within and beyond their cultural contexts that encompass several cultural aspects, such as values, identity, traditions, communication and cultural heritage (Huffling and Stevenson, 2019; Vietze et al., 2019; Milner, 2021). Furthermore, to 'enhance a more equitable and social justice-oriented science pedagogy', Upadhyay (2023, p.8) has introduced the culturally responsive science pedagogy (CRSP) framework in the 'enactment of the science curriculum, pedagogy and assessment for Asian countries'.

Our proposed Culturally Responsive Methods (CRM) connect cultural heritage to project work which enables learners to be part of the cultural heritage of others. We thus interpret cultural heritage in a broader context (UNESCO Institute for Statistics, 2009), which includes both the tangible and intangible elements, i.e., the historical and artificial objects are connected with cultural values, in the sense that objects may interact with memory and evoke certain values detached from material aspects (Vecco, 2010). Bridging students' backgrounds with historical and regional stories, buildings, pictures, etc. thereby activates links to positive personal ideals and historical culture and personal backgrounds including different prior knowledge and variable learning types. These CRM with direct adaptation to cultural heritage connect to a part of the learners' environment allowing for individual support of diverse learners in a way that addresses the individual background, increases the interest in the taught subject and helps to reduce the obstacles that come up when learners have low prior knowledge, misconceptions, or insufficient interest. The students could find a way to bridge the scientific content with their own and others' lives thereby realizing the importance of the content and increasing interest in learning, improving self-assessment and stimulating creativity.

Furthermore, various areas and approaches to learning and interests are addressed

in a very natural way that allows self-consistent inclusion of interdisciplinary and international approaches as addressed in actual developments of novel interdisciplinary and international study programs. Offering such methods that bridge to learners' personal experiences enables them to take an improved responsibility for their learning while bringing the past with the present and future knowledge and experience thereby not only strengthening the link to prior knowledge using individual images, examples, experiences, etc. but also developing an increased responsibility for the environment in general (including actual concerns, beliefs, problems). Finally, this approach may help to identify and preserve cultural and natural heritage for future generations in the spirit of international and interdisciplinary cooperation.

Methodology

Research Design

Our study aims to develop Culturally Responsive Methods in mathematics teaching and learning that dynamically adapt to the student's cultural heritage and improve the context-specific inquiry processes and project-based teaching (Scherff & Spector, 2011; Mack, Winter & Soto, 2019). Based on this aim, we formulated the following research questions:

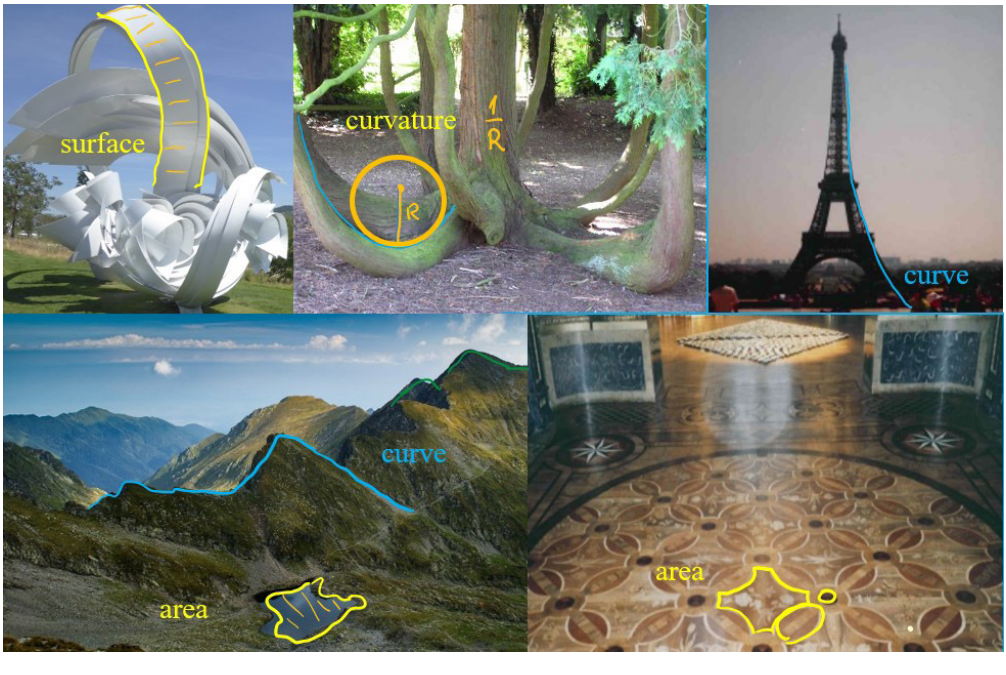
- How does the CRM support the creative learning process of undergraduate students?
- How does the CRM influence undergraduate students' learning of mathematics concepts and self-assessment?

A design-based research (Anderson & Shattuck, 2012; Barab & Squire, 2004) was the approach used to understand real-world practice, involving multiple dependent variables. In addition, the participants are treated as co-participants in the design. We chose project-based work as a form to implement the CRM with embedded cultural heritage and to collect data since it offers a quite flexible way to dynamically adapt to the student's cultural heritage and supplement traditional methods within the frame of self-chosen tasks. The project work was combined with the traditional written test at the end of the semester. Both parts, i.e. the written test and the project work, contributed 50 per cent to the overall grade. With a comparison of the results of the written test only with the grade resulting from the combination of the written test and the project work (50% of the final grade respectively), we explored whether the culturally responsive methods based on cultural heritage are associated with better learning outcomes.

In the first stage of the designed-based research, following the lectures' learning objectives, we started developing Culturally Responsive Methods connecting various scientific concepts with cultural heritage and with individual students'

projects. Students are offered multiple opportunities and topics by predefined lists that may be extended with additional topics by students. For a given lecture, this type of list includes appropriate links to cultural and historical topics, typical national/regional issues, daily life- issues, and others. Given examples may inspire students to find more and to select topics they like. Students were thus given the task of finding their examples appropriate to the project and concepts based on cultural heritage. This process allows each student to find a method that is optimally fitting to their preferences for memorization and own interests, for example, computer simulations for visualizations, self-developed sample tasks, drawings, figures, personal situations, and stories. The topic and method were chosen in a way that the student feels supported in the effort to segment the material into manageable and meaningful chunks and to link those with personal meaning, for example, in terms of personal or global heritage, personal or global/internationally relevant interest, and so forth. Guidelines for this project work and opportunities for dynamic feedback and support were given throughout the whole project. The CRM were used in Differential Geometry, Variational Approach, and Mechanics courses and the topics of these courses can quite easily be transferred to real-life scenarios and help students find their "own examples". However, we would like to point out that links to their own lives can similarly be found in any other area.

Figure 1. Examples of curves and shapes in individual cultural heritage - challenging to describe and beautiful to remember



As a first example, we refer to the 'Differential Geometry' lecture course and focus on curves and areas in space topic. Project work may focus on the description of curves and/or surfaces of real shaped objects (selected by the students, e.g. having a particular meaning for the corresponding author of the project work). Regional and national aspects may come into play by creating curves and surfaces describing nature (fields, mountains) and historical buildings (leading to questions concerning e.g., curvature and torsion but also connections to forces and stability and the adjustments that are required when lines and surfaces are not only geometrical objects but additionally describe real mass distributions). All these may be explored and discussed by suitable mathematical approaches, for example, finding suitable parametrized representations of curves and surfaces (Figure 1), using interpolation for interpretation of relics and ancient ruins, relating bending curves and surfaces to explore material fatigue and stress in technical examples and so forth. The lecturer may thereby give recommendations depending on an assessment of learning types, he/she may also show selective examples to reinforce the development of own ideas.

The basic idea of the course Variational Calculus is the challenge to find functions that optimise the values of quantities that depend on those functions. Examples of this can be found almost everywhere in historical artefacts and nature. Practically, all fundamental laws of nature can be derived from such principles of variation: many shapes in the environment arise because matter tries to find an optimum in potential energy while experiencing forces that describe the cohesion of matter, elasticity, gravity etc. One can think, for example, of the problem of determining the shape of a hanging chain suspended at both ends. The mathematical solution is a function (in mathematics called 'catenary') that minimises the variational potential energy of the chain. Further examples are the minimum time for a light path (Fermat's principle) or the maximum stability of a bridge, an arch or a tower in a historic building. Finally, the content of the mechanics course can be applied easily to a large variety of examples of cultural heritage and environments. Those range from the continuum mechanics of natural matter and historical buildings to the rigid body dynamics of technical inventions, that can be explored with mathematical methods.

Sample

A purposive sample of 70 undergraduate students was selected for the first phase of our research. Students were in the first and third years of their study program and took the courses mentioned above in 2021. For organizational reasons and purposes of the research, the questionnaire was done with a purposive sample of 30 students and for the Certainty-Based Marking (CBM) analysis with a sample of 40 students.

Methods to Collect and Analyse Data

For the first phase of our design-based research, data was collected using project-based work, questionnaires, written tests and self-evaluations in combination with quizzes. The quantitative data collected in the first phase of the project was analysed using descriptive and inferential statistics. For the following phase of our research, additional data will be collected using maps of concepts, stories and reflective journals, which will be analysed using content analysis.

Table 1 Six Levels to Promote Creativity in Higher Education (based on Jahnke & Haertel, 2010)

Level	Description
6 Original, completely new ideas	Cannot be forced, if necessary, promotion of many ideas through creativity techniques and suitable environments, allow/promote mistakes
5 Promoting a new culture of thinking	Change of perspective, breaking through habitual patterns and routines, adopting a different attitude, reflection on one ‘s own creativity and thinking structure, knowledge about how the brain works
4 Promotion of creative learning	Creating/providing ideas, networking, transfer
3 Increase the motivation to learn: promote flow	promote flow, e.g. with metaphors, humour, variety, enthusiasm, develop interesting questions/problems, establish practical relevance
2 Promotion of independent working	Acquiring knowledge yourself, controlling learning processes independently, making your own decisions
1 Promotion of independent, reflective learning	Acquire knowledge yourself instead of just taking it over, engage in inner dialogue, think outside the box, encourage to question the familiar

Questionnaire based on Jahnke & Hartel’s (2010) socio-construction creative model

To investigate the efficiency of CRM based on project-based work and personal cultural heritage on the social construction of creativity, we collected data using a questionnaire (Table 1) based on Jahnke and Haertel (2010). According to Jahnke and Haertel (2010), creativity leads to flexible and critical thinking, the generation of problem-solving strategies and thus efficient learning. This questionnaire addresses six levels to promote creativity and provides descriptions concerning their meaning in a learning process. In particular, the addressing of multiple levels is decisive for an efficient improvement in learning. In a previous publication by one of the authors (Gehrig et al., 2022) this questionnaire was already applied to a portfolio work combined with stepwise feedback to students.

Certainty-Based Marking

A very suitable scheme to evaluate how realistic are students' self-assessments of their performance is 'certainty-based marking' (CBM) which is a method that maps and evaluates both correctness and self-assessment in one scheme (Yuen-Reed& Reed, 2015). Thereby realistic self-assessment is rewarded leading to an increased reflection. The principle is as follows: firstly, the students have to answer a question about the content of the study course and secondly, they have to mark how confident they are that their answer is correct.

Table 2 Certainty-based marking (Yuen-Reed and Reed, 2015; Kanzinger and Gehrig, 2022)

Marking scheme	Degree of uncertainty	
	certain	uncertain
correct answer	5	4
wrong answer	0	2

As can be seen in Table 2, most points are received in the answer is correct and the student is sure about that. In case the student is not so sure, he/she receives 4 points. A wrong answer is still given 2 points if the student realizes that he/she is not sure – a reward for a realistic self-assessment. On the other hand, feeling sure while having given a wrong answer leads to zero points – in agreement with the danger a wrong answer given with strong conviction means in real life! To analyze the empirical data and to test if there is a correlation between knowledge and self-assessment we conducted statistical tests on the students' results (i.e. percentage of received CBM points vs percentage of correctness in their answers).

Results and Analysis

Social Construction of Creativity and Learning

The questionnaire was completed by 30 students (2/3 in the first year, 1/3 in the third year). All students (100%) estimated their learning level and knowledge as greater, more firmly anchored and deeper when the projects were included compared to learning for an oral or written exam only (i.e. without consideration of project work as part of the final grade of the study course (level 1). Support in independent reflective learning was perceived by 75% of the students (level 2). An encouragement of self-organised learning by developing approaches and exploring the environment and cultural background with topics related to the lecture content was confirmed by 85% of the students (level 3). An increase in motivation by working creatively with their individual stories was stated by 80 % of the students. They reported that they enjoyed developing and implementing the projects (level 4). A promotion of a new culture of thinking was reported by

85% of the students. In particular, they reported that they gained competence and experience and felt encouraged to apply their knowledge to personal situations, historical backgrounds and real-world situations (level 5). The ability to develop entirely new discovered ways to explore things was perceived by 80 % of the students (level 6).

Generally, the results confirm that CRM with direct adaption to individual cultural heritage, backgrounds and experience demonstrate a concrete improvement on all creativity levels. The improvement in levels 1—3 is high but not as high as in a previous study (Gehrig et al., 2022). This result is because we had less feedback included in our present study. The professor was also available for questions and discussions, but the feedback was not as much part of the method as in the previous study indicating that feedback methods probably reinforce contributions in the first three levels. It is noticeable and important, however, that the ability to develop entirely newly discovered ways to explore things was perceived by 80% (a significantly larger value than in the previous study which was 70%). We can conclude that the coupling to students' cultural heritage is a strong factor qualifying them to apply their knowledge to completely new situations. It is this competence which is of particular importance for professional life. The results are therefore promising and motivate the further development of such methods.

Students Self-Assessment: Certainty-Based Marking (CBM)

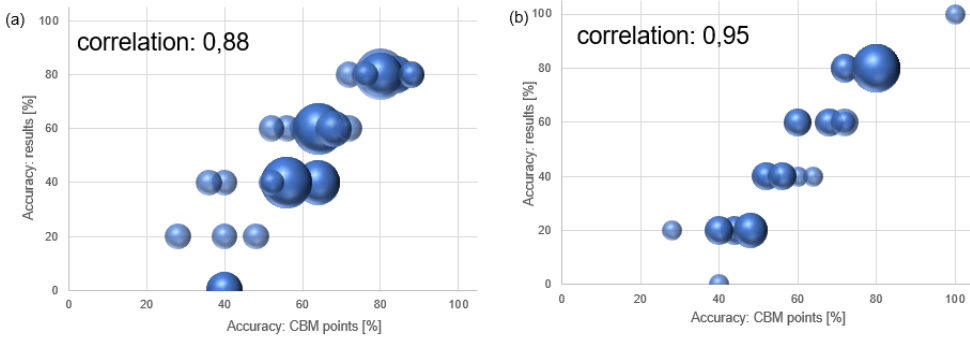
In order to gain information on students' self-assessment with respect to the learning outcome, we applied the Certainty-Based Marking model to students of a study course (40 students in the third year) where the students submitted a project work in the first part of the semester (Project 1) and a work in the second part of the semester (Project 2). They were afterwards additionally asked to answer small comprehension questions and to self-assess their learning using Certainty-Based Marking as a formative assessment of learning status and as a method to initiate self-regulated learning. The results are shown in Figure 2. We evaluated what proportion of the work was correct in each case (y-axis) and what proportion of the maximum CBM points was achieved (x-axis). By asking students to assess themselves, they become aware of any knowledge gaps that may still exist and increase their learning effort (e.g. 40% Accuracy in self-assessment and 0% in results means, that the students in that group realize by 40% that their work is wrong and that they have to do more).

A value near the diagonal thus signifies that a student has a good estimation of how well he/she performed (see also area of learning (1) in Figure 3(b)). Thereby, low values mean that he/she is not good but is aware of the situation. High values (near 100%) mean that he/she performed well and has a reasonable self-assessment of her work (see also area of learning (2), Figure 3(b)). The

values on both sides of the diagonal show misconceptions (values closer to the x-axis, area of learning (3) in Figure 3(b)), i.e. the students think they are better than they are, and uncertainty (values closer to the y-axis, area of learning (4) in Figure 3(b)), i.e. students performed reasonably well but do not trust their knowledge sufficiently.

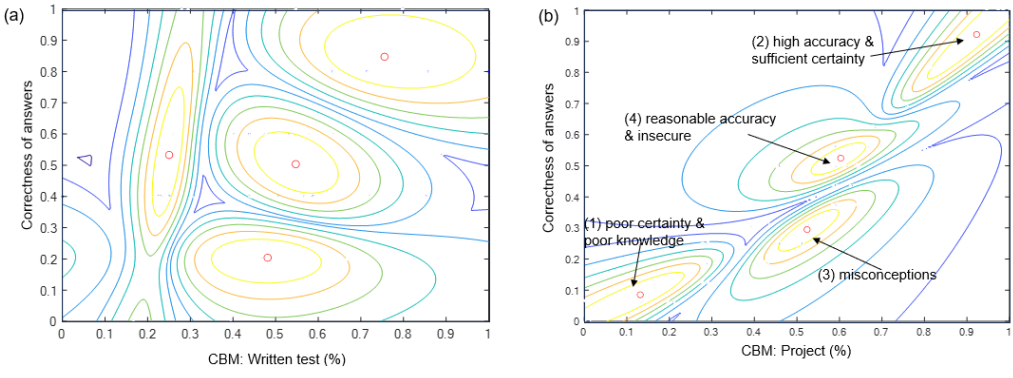
The dependence of the final grade (Accuracy in the results, y-axis) on the Accuracy of the self-assessment (CBM points, x-axis) clearly shows that the competence of the students for realistic self-assessment is strongly correlated to the real output. In particular, the correlation is stronger in the situation displayed in Figure 2(b), i.e. the students get better in a realistic self-assessment when they have already finished the second project, i.e. are more experienced and profited from the experience of the projects.

Figure 2. Correlation between obtained results and self-assessment (CBM)



To further investigate the influence of project work, we additionally performed a cluster analysis using Matlab. We compared a data set for a grout where the final grade was given based on one written test result at the end of the semester (Figure 3(a)) with a comparable group (same size and average performance) where project work was included during the semester (Figure 3 (b)). The number of four clusters was predefined, representing four typical areas of learning: (1) poor certainty &poor knowledge, (2) reasonable accuracy and insecure, (3) misconceptions, (4) high accuracy and sufficient certainty. The algorithm groups the data set, the four points mark the centers of the subsets. The self-assessment can be used to extract both learning types and the need for support. In particular when working on own examples (with a context to personal situation and background as extended cultural heritage) the percentage of the area of misconceptions and insecure work was reduced and the percentage of high accuracy in combination with certainty in finding the solution increased.

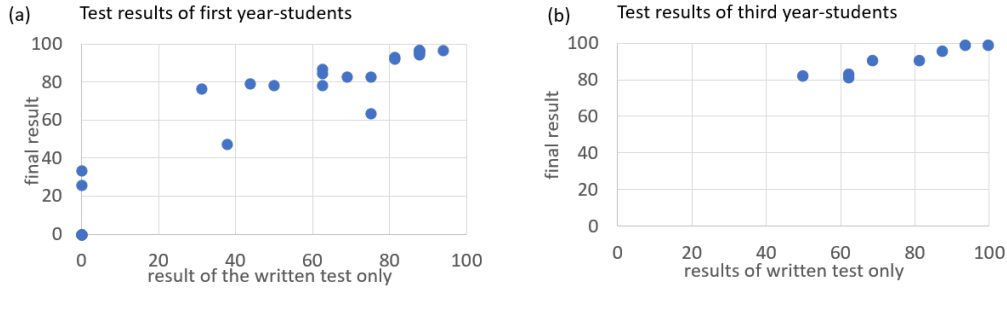
Figure 3. Identification of four areas of learning with cluster analysis



Learning Experiences

The results of the self-assessment discussed in the previous section suggest that learning is increased through the integration of self-chosen projects based on extended CRM cultural heritage. Integrating the result of the project as part of the final grade of the course (here: 50%) improves both, the ability to assess yourself well and the quality of the final result and thus the students' knowledge. In the final step, we aimed to find out if the learning improvements obtained by CRM depend on students' prior experiences. We compared results for first-year and third-year students (Figures 4 and 5). In Figure 4, the x-axis shows the grade (in percentage, i.e. 100% means „everything correct") of the written part only. The y-axis was the entire results composed of the test (50% written + 50% project).

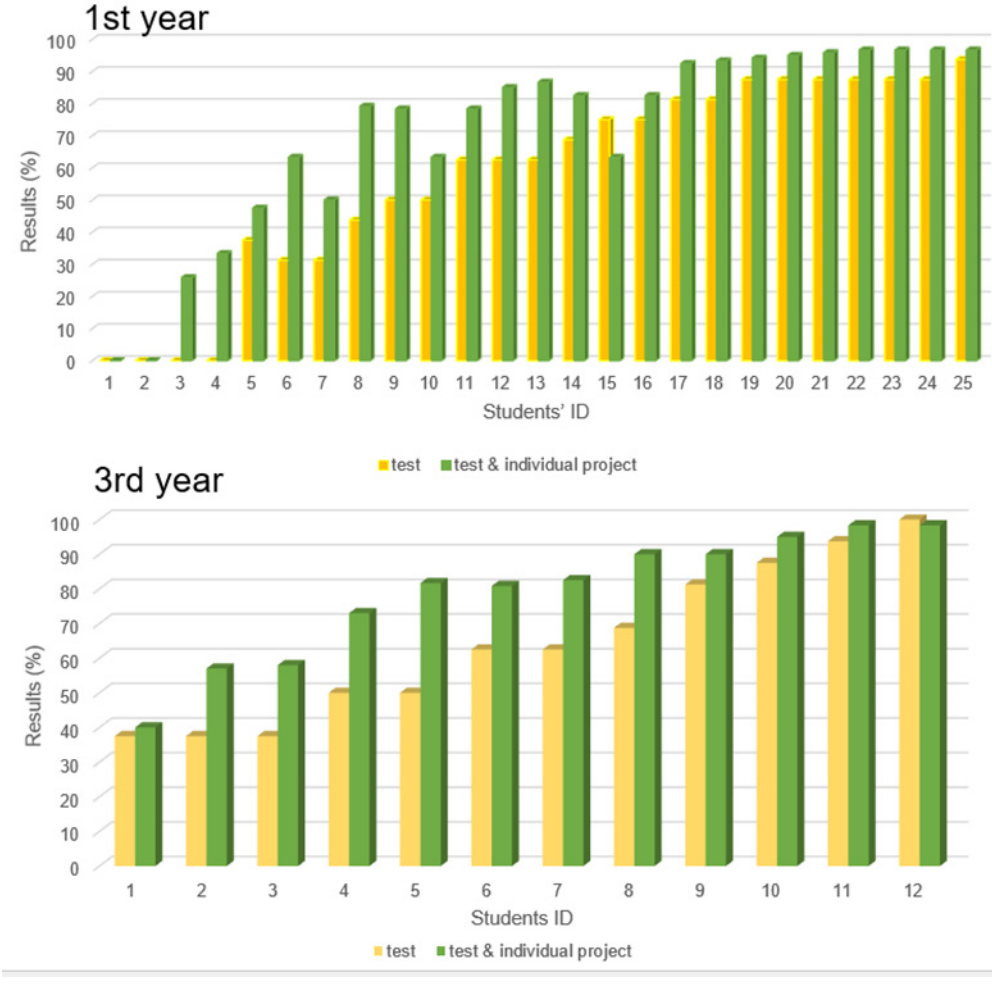
Figure 4 Comparison of test results for first-year and third-year students



The majority of students' learning improved through the additional project work (first year and third year). The group of very weak students (three, with 0 points

in the written exam) could not be saved via the project work. Nevertheless, they did some work and might succeed in the second trial. The very good students (five, with more than 80% in the written exam) have similar results, the results are simply shifted to better grades for the project work. The students in the regime between 40-80% in the results of the written test only typically received much better results in the final test (for example 50% in the written test and 80% in the project work).

Figure 5. Comparison of first-year and third-year students tests results (individual results)



As expected the observed effect is particularly important in the first year where students coming from a different background have to arrive in the university system and way of (mathematical) thinking. Figure 5 further confirms this behaviour by showing the results of the test only as well as the test and individual project work for the individual students. While the integration of the projects in

most cases generally increases the learning outcome the results for the first-year students particularly underline the importance of individual project work, i.e. a considerable proportion of students with large differences in the results marking the need and pathing the way to an improved learning in future.

Conclusions

In conclusion, our first results demonstrate that integrating culturally responsive methods with direct adaptation to cultural heritage, history, technological applications, and personal experiences, helped undergraduate students to create sustainable learning and creativity. In the long-term planning, the collections may lead to a kind of dynamic exhibition representing individual courses, parts of which may be didactically arranged on internet platforms to reach an even broader audience thereby sensitizing for intercorrelations between math, science, and cultural heritage. In the second phase of the project, by using additional culturally responsive methods, i.e., an overall map of concepts, relations, ideas, stories, applications, and experiences, in addition to students' explanations and reflections, we would like to investigate how these culturally responsive methods had help students to find themselves as a responsible part of a whole and also to get to know the elements represented by others, including the interpretation and meaning provided by the people behind the remaining project ideas. So far, the CRM have been applied to study courses mainly attended by German students, with a minority of foreign students. For future work, we plan to apply CRM to international study programs. Having this new variable might be particularly challenging due to a larger difference in personal background and knowledge. However, including this variable, will allow us to gain deeper insight into how to strengthen learning and remembering in heterogeneous groups.

References

Anderson, T., & Shattuck, J. (2012). Design-Based Research: A Decade of Progress in Education Research. *Educational Researcher*, 41(1), 16-25. https://doi.org/10.3102/0013189X11428813

Averill, R., Anderson, D., Easton, H., Te Maro, P., Smith, D. & Hynds, A. (2009). Culturally Responsive Teaching of Mathematics: Three Models from Linked Studies. *Journal for Research in Mathematics Education*, 40(2), 157-186.

Barab, S. & Squire, K. (2004) Design-Based Research: Putting a Stake in the Ground. *The Journal of the Learning Sciences*, 13(1), 1-14.

Cunningham, E. (2001). Culturally Responsive Pedagogy. *Counterpoints* 130, 85-102.

Dodo Seriki, V. (2018). Advancing alternate tools: Why science education needs CRP and CRT. *Cultural Studies of Science Education*, 13(1), 93-100. https://doi. org/ 10. 1007/ s11422- 016- 9775-z

Gay, G. (2000). *Culturally responsive teaching: Theory, research, and practice*. (2nd

ed.). Teachers College Press.

Gehrig, E., Kanzinger, A., & Zender, J. (2022). Portfolios as Formative Assessment in Applied Mathematics: Influences of Portfolios on the Teaching-learning Process. In Shelley, M. & Sahin, I. (Eds.), *Studies on Education, Science, and Technology 2021* (pp. 33-58). ISTES Organization.

Huffling, L. & Stevenson, A. (2019). Creating Culturally Relevant Science Stories. *Science Scope*, 42 (9), 68-77.

Jahnke, I., and Haertel, T. (2010). Kreativitätsförderung in hochschulen–ein rahmenkonzept. *Das Hochschulwesen*, 58(3), 88-96.

Kanzinger, A. & Gehrig, E. (2022) *Certainty-based-marking - eine kompetenzorientierte Prüfungsmethode,* 56th Annual Meeting of Society for Didactics of Mathematics

Ladson-Billings, G. (1995). But that's just good teaching! The case for culturally relevant pedagogy. *Theory into Practice,* 34(3), 159-165. https://doi.org/10.1080/00405849509543675

Mack, K. M., Winter, K., & Soto, M. (Eds.). (2019). *Culturally responsive strategies for reforming STEM higher education: turning the tides on inequity*. Emerald Publishing Limited.

Milner, H.R., (2021). *Start where you are, but don't stay there: Understanding diversity, opportunity gaps, and teaching in today's classrooms*. Harvard Education Press.

Ratner, C. (2008). *Cultural Psychology, Cross-cultural Psychology, and Indigenous Psychology*. Nova Science Publishers.

Scherff, L., & Spector, K. (2011). *Culturally relevant pedagogy: clashes and confrontations*. Rowman & Littlefield Education.

Tate, W. (1995). Returning to the Root: A Culturally Relevant Approach to Mathematics Pedagogy, *Theory Into Practice*, 34(3), 166-173.

UNESCO Institute for Statistics (2009). *Cultural heritage. Definition*. Retrieved January 22, 2024, from https://uis.unesco.org/en/glossary-term/cultural-heritage.

Upadhyay, B. (2023). Culturally Responsive Science Pedagogy for Social Change and Personal Transformation. In Halim, L., Ramli, M. & Ali, M. (Eds), *Culturally Responsive Science Pedagogy in Asia- Status and Challenges for Malaysia, Indonesia and Japan* (p.8), Routledge.

Vecco, M. (2010). A definition of cultural heritage: From the tangible to the intangible. *Journal of Cultural Heritage, 11*, 321–324. doi.org/10.1016/j.culher.2010.01.006

Vietze, J., Juang, L. & Schachner, M. (2019). Peer cultural socialisation: a resource for minority students' cultural identity, life satisfaction, and school values. Intercultural Educatıon. 30 (5), 579–598, https://doi.org/10.1080/14675986.2019.1586213.

Yuen-Reed, G. & Reed, K. (2015). Engineering Student Self-Assessment Through Confidence-Based Scoring. *Advances in Engineering Education*. 4(4), 1-23.

Understanding Professional Learning in Higher Engineering Education through Signature Pedagogies

Mijung Kim[1], P. Janelle McFeetors[1], Qingna Jin[2], Kerry Rose[1]

[1]University of Alberta, Canada

[2]Cape Breton University, Canada

Increasing numbers of studies in STEM fields indicate there are higher gains in student outcomes in classroom environments that implement evidence-based teaching approaches focused on learners and their learning, when compared to traditional lecture-based classrooms. However, despite the compelling results, studies show that lecture is still the most common instructor behavior in undergraduate education. Professional learning programs have been shown to be crucial in closing this gap and promoting sustained adoption of student-centred teaching approaches in undergraduate education. In order to understand how faculty members in engineering experience the development of their pedagogical understandings and practice, we invited seven engineering professors to participate in a professional learning program, Scholarship of Pedagogy and Application of Research Knowledge in Engineering (SPARK-ENG), designed specifically for engineering education. The participants' interactions during the community of practice and individual and group interviews after completing their modules were recorded, transcribed, and analyzed through a thematic analysis process. The study findings indicated that the participants experienced the complexity of pedagogical understanding and decision-making over knowledge and practice of engineering, and valued community-based interactions that enabled reflection on their teaching. The study points to the need for further research and discussion surrounding professors' perspectives of the nature of engineering knowledge, and pedagogies that might be congruent with these.

Keywords: Signature pedagogies, Higher education, Engineering professors

Research Background

Recent calls in higher education have specifically noted the need for faculty to consider a change from lecture-based teaching towards a more evidence-based, student-centred approach. These calls are supported by the increasing number of studies that indicate overall gains in student outcomes in environments that implement student-centred pedagogical approaches, as compared to traditional lecture-based classrooms (e.g., Dębiec, 2017; Derting & Ebert-May, 2010). However, despite these compelling results, studies have also shown that, at

least in engineering undergraduate education, lecturing is still the most common instructor behaviour; in a recent survey of engineering educators teaching in Canadian universities, about 95% of instructors responded that the main instructional mode that they use is lecture. Although Canadian engineering educators claim to place a strong emphasis on their teaching, student reports of experiences other than lecture-based instruction are also limited with respect to their undergraduate engineering education (Nelson & Brennan, 2019). This study's aim was to examine how engineering faculty members experience their learning about student-centred pedagogies through an online, module-based professional learning program called Scholarship of Pedagogy and Application of Research Knowledge in Engineering (SPARK-ENG).

Conceptual Theoretical Framework

Teacher Understanding of Pedagogy: Signature Pedagogies

At universities, many professional development (PD) programs are designed for instructors across disciplines and provide teaching strategies and approaches for generalized contexts; but, university lecturers have expressed that they prefer more discipline-specific PD programs (MacPhail et al., 2019). Shulman (2005) proposed that instructors' understanding of pedagogies need to be developed in the context of specific professions and that these unique context-dependent teaching strategies could be termed, 'signature pedagogies.' Signature pedagogies include the characteristics of teaching and learning of specific disciplinary knowledge, skills, and cultural norms within professions. Professional knowledge and performance are often discipline-specific and situated in the culture of professions: to teach students to become professionals in their own fields includes knowledge development, but also ways of acting, performing, and practicing as professionals in their fields. Therefore, PD programs incorporating signature pedagogies in engineering have the potential to effectively help educators prepare students for the multifaceted challenges they will encounter in their future careers as engineers.

Situated Learning

Situated learning theory (Lave & Wenger, 1991) postulates that learning occurs as a result of socially and culturally embedded interactions and relationships and is thus situationally (socially and culturally) dependent. This is congruent with the theoretical conceptualization of signature pedagogies within pedagogical knowledge domain as learning is seen as a social process of enculturation. In this study, we draw on situated learning as a "a process of interaction and relationship around a specific domain and which occurs within a social, cultural, and historical context, resulting in spontaneous learning" (Ebbers, 2015, p. 650). Learning contexts often exist within specialized communities of practice (CoP)

where both novices and experts collaborate in genuine practices unique to the domain (Ebbers, 2015). The context that the participants are situated within in this case includes the disciplinary consideration by the participants of what an engineer is and what one does. In this study, we considered the uniquely situated context of engineering education as foundational to our understanding of how participants interacted with the module-based learning program and each other as part of a CoP.

Research Processes

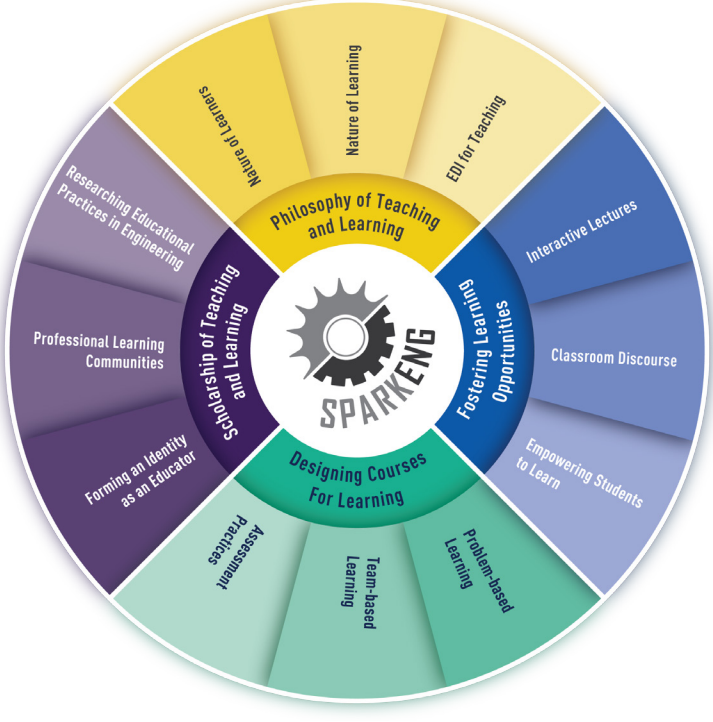
Procedure

At the University of Alberta (Canada), the Centre for Mathematics, Science, and Technology Education partnered with the Faculty of Engineering to design a professional learning program that would support engineering professors to enhance their pedagogy. The aim of the program was to develop a culture of teaching that supported active, student-centred practices grounded in evidence-based approaches. The 4 themes that cluster 12 modules was designed collaboratively, based on mathematics and science educators' background in pre-service teacher education and research programmes, along with Faculty of Engineering's recognition of areas of growth. After identifying the focus of each module in SPARK-ENG (see Figure 1), content writers were asked to develop engineering-specific materials. This design element was critical because of the importance of a preference for discipline-specific professional learning and because learning is situationally dependent. In particular, content writers selected case studies that were about teaching in undergraduate engineering classrooms. They interviewed both engineering professors and engineers and incorporated podcast interview clips with the growth of future professional engineers in mind. The required readings were written by engineering educators. As well, workplace learning task suggestions were tailored to be engineering content specific. At times when required readings were difficult to locate, writers drew on post-secondary disciplines such as physics education and mathematics education as these are foundational content areas in first year engineering courses.

After two modules, the Nature of Learners and Problem-based Learning, were complete, we piloted them with a small group of engineering faculty. A pilot was important to determine, early in the module writing process, what immediate aspects could be improved for a better learning experience, guiding the modification of the pilot modules and for the overall organization of the whole program later. The pilot participants engaged with one module per month. In the first week, they reflected on their previous experiences, considered a provoking case study (short video or podcast), and read two articles (a short one focused on a professional audience and a longer one focused on a scholarly audience). In the second week, they learned about approaches and considerations

in developing their pedagogical practices through reading content, interpreting infographics, listening to podcast clips, and viewing curated videos. During this week, participants were asked to apply what they were learning to create a workplace learning task: this is an evidence-based practice that they could try in their classroom given supports and templates from the module (e.g., create and implement a student survey). The first two weeks were completed asynchronously online. In the third week, they gathered synchronously as a group – called a Community of Practice – to discuss what they learned from the readings/content and the successes and challenges of trying a new teaching approach for their workplace learning task. Completing a module took approximately 10 hours over the month.

Figure 1. SPARK-ENG program overview.



Data Collection & Analysis

We used a qualitative case study approach (Merriam, 1998). We inquired into how engineering faculty members learned, reflected, and developed their understandings of pedagogies of engineering teaching through the SPARK-ENG pilot modules: 1) The Nature of Learners (NL) and 2) Problem-based Learning (PBL). We had 7 engineering faculty members participate in the pilot project. They had a range of teaching experience, from less than 1 year to more than 10 years. They engaged in all components of the two modules.

We collected qualitative data from two video recordings of Communities of Practice sessions, one video recording of a one-hour focus group, and video recordings of individual, semi-structured interviews at the end of the pilot project with each participant. The data were subjected to qualitative thematic analysis (Braun & Clarke, 2006). Thematic analysis offers researchers flexible guidelines for the data coding process where the aim is to identify themes and patterns that explicate interpretive understandings of participants' experiences. Specifically, strategies of open, axial, and selective coding were employed. To develop the trustworthiness of the data analysis, cross-checking among members was conducted.

Findings

Practical Approaches to Teaching Engineering

In this situated learning context, namely that of undergraduate engineering professors, we noticed the participants were willing and receptive toward learning to enhance teaching. It was clear, though, that the participants were more receptive to module components that were seen to have a more practical approach in the two modules presented, Problem-based Learning (PBL) and the Nature of Learners (NL). In general, they found that PBL was more practical overall and thus seen to be more relevant and helpful for them as instructors. For example, a participant mentioned, "To be honest, the PBL module ...matched my expectations. I learned something very practical. The other one [NL], I learned the concept but still I feel it's too theoretical right now." However, for these participants, learning about their learners (i.e., who their students are) was not seen to be as critical for improving their teaching competency compared to learning some strategies for teaching classes. The participants often expressed how learning about their students was considered to be less relevant, less effective, or less possible to implement in their particular classroom context. For example, one participant shared:

> Looking after different students actually is not easy, especially under the North American system. … I do not see any chance to implement such assistance because, for a typical class with 50 to 100 students, the professor has no chance to look after every aspect of the students.

Understanding the context of class size helped us address this challenge in the subsequent module writing. Out of the two modules piloted, the participants valued the implementable practicality of PBL with whole and larger classes as a helpful and effective way to develop their pedagogy and teaching practice. Although they acknowledged that knowing their learners could allow for better instruction, most participants reported that they had little experience and saw few opportunities to do this with large undergraduate classes, and thus they

considered the NL module to be more conceptual in terms of pedagogy and thus not as useful.

Influence of Participation in Practices to Become an Engineer

With respect to pedagogical approaches, our participants expressed that some student-centred strategies and ideas were seen to be not typical of traditional engineering education culture. This was illustrated by responses that saw student-centred approaches as outside of the common teaching repertoire of undergraduate engineering education practice. Participants explained how they valued understanding their students, but only with respect to how this affected the students' ability to grasp content knowledge and skills that they are teaching. When the module content was more congruent with the participants' view of the Nature of Engineering and how students could work toward becoming part of this community, this was seen as instructors helping their students to engage in the application of knowledge to new problems. For example, a participant explained:

> [As engineers] We prefer things more practical … We learn things to solve real-life problems. Problem-based learning helps to bridge the gap between classroom and the real world. ... We need to find out how can we help students realize that they not only need to understand the information that we're teaching but that they need to be able to apply the information that we are teaching.

Participants also viewed this as supporting students' development of an engineering identity, as engineers work to solve clients' problems by drawing on disciplinary knowledge. Participants often expressed this as they spoke of what types of pedagogical approaches resonated with them, especially as they learned to become engineers.

Importance of Engagement with the Community of Practice (CoP)

The participants clearly appreciated their engagement in the Community of Practice (CoP) discussions. During the CoP discussion after individual completion of each module, the participants reflected on teaching and learning with colleagues and shared their experiences, insights, wonders, and struggles for classroom implications of the module contents. They were able to clarify interpretations of the readings and set them in the context of approaches tried in their classrooms. When they shared how they implemented certain pedagogical strategies in their own classrooms and how these worked (or not), they reported that they became more aware of the potential and feasibility of classroom implementation, which often encouraged them to accept these strategies as evidence-based, effective, and trustworthy teaching pedagogies for engineering classrooms.

> During the CoP you can really communicate with others. ... They can give [their] feedback, like they said, 'we have already done that and it's very, very effective.' Those things ... [are] hard to learn from the materials, ... [they are] very personal experiences. ... you hear that someone in your university, they use this and it's useful. That's something you want to hear.

Consistent with the Nature of Engineering, participants would also share practical solutions to challenges colleagues' experienced when trying a new pedagogical strategy. Another encouraging result is that participants began to develop new perspectives as they learned from each other. Thus, the participants viewed their engagement in the CoPs as an essential element of their learning as they developed further understanding, critically examining ideas while situated within a learning community of discipline-related colleagues.

Discussion and Implications

There are various influences that inform teachers' pedagogical decisions in classrooms. External factors such as time constraints, prescribed curriculum, classroom environments, etc. influence what and how to teach. Yet, research also explains that teachers' epistemic beliefs in disciplines deeply influence their teaching approach. In this study, the participants also highlighted their concerns around constraints of time and class sizes toward certain pedagogical approaches. At the same time, it was also evident that the participants' perceptions and beliefs on what engineering and teaching engineering meant to them seemed to influence what to prioritize for effective teaching throughout the modules. The participants often suggested that engineers learn by doing and/or problem-solving, and that engineers enjoy practical applications of ideas. Thus, they believed that their students, who chose engineering as a major, valued similar ways of learning. Their perceptions about engineering reflected descriptions of engineering as a discipline of practical, real life knowledge and problem solving and this transferred to their preferences for pedagogical strategies in their own disciplines. From this study, we suggest that there is a need for PD programs in higher education that are specifically aligned and developed within and for their own disciplines, that is, the approach of Signature Pedagogies.

Acknowledgements

This research was funded by the Centre for Mathematics, Science, and Technology Education (Faculty Member Funding), the Faculty of Engineering (Grant; NSERC Undergraduate Student Research Award), the Faculty of Education (Roger S. Smith Undergraduate Student Research Award; Support for the Advancement of Scholarship Grant), and the University of Alberta (Supplementary Research Assistance Funding). We are grateful to the participants who shared their

perceptions of their experiences in SPARK-ENG modules with us.

References

Braun, V., & Clarke, V. (2003). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Dębiec, P. (2017). Effective learner-centered approach for teaching an introductory digital systems course. *IEEE Transactions on Education*, *61*(1), 38-45. https://doi.org/10.1109/te.2017.2729498

Derting, T. L., & Ebert-May, D. (2010). Learner-centered inquiry in undergraduate biology: positive relationships with long-term student achievement. *CBE—Life Sciences Education*, *9*(4), 462-472. https://doi.org/10.1187/cbe.10-02-0011

Ebbers, S. (2015). Situated learning. In J. M. Spector (Ed.), *The SAGE encyclopedia of educational technology* (pp. 649-651). SAGE.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge University Press.

MacPhail, A., Ulvik, M., Guberman, A., Czerniawski, G., Oolbekkink-Marchand, H., & Bain, Y. (2019). The professional development of higher education based teacher educators: Needs and realities. *Professional Development in Education, 45*(5), 848-861. https://doi.org/10.1080/19415257.2018.1529610

Merriam, S. B. (1998). Case study as qualitative research. In S. B. Merriam (Ed.), *Qualitative research and case study application in education* (pp. 26-43). Joseey-Bass.

Nelson, N., & Brennan, R. (2019). Effective learning environments: Is there alignment between the ideal, the actual, and the students' perspective? *Proceedings of the Canadian Engineering Education Association (CEEA), 19*(paper 109), 1-7. https://doi.org/10.24908/pceea.vi0.13787

Shulman, L. S. (2005). Signature pedagogies in the professions. *Daedalus, 134*(3), 52–59. https://doi.org/10.1162/0011526054622015

Navigating The STEM University Transition: Effects of a Combined Mindset and Learning Strategy Online Intervention

Malte Diederich[1], Jana Rehberg[2], Verena Spatz[1], Thomas Wilhelm[2]

[1]Technical University of Darmstadt, Germany

[2]Goethe University Frankfurt, Germany

Mindsets (implicit theories of intelligence) can have an impact on learning and play a particularly important role when facing challenges such as transition from school to university. This study looks at the effects of a single online intervention on students' motivation and intentions to change major or drop out. The intervention combines a mindset message with learning strategies adapted to the university STEM context in a symbiotic way. With a questionnaire survey (T1 at the start of the semester, T2 two months later), domain-general mindset, domain-specific mindset and domain specific academic self-concept are measured for an intervention group (IG, n=89) and a baseline group (BL, n=278). Results suggest a positive impact of the intervention on mindset and self-concept. Regarding intentions to change majors or drop out (measured only at T2), we report in general low intentions among the participants. Only a subgroup that reported continuous struggle during the semester showed higher intentions, which were lower amongst the students who participated in the intervention.

Keywords: STEM university studies, growth mindset, intervention.

Introduction

STEM subjects play an important role in many problems of this century and there is a high demand for professionals with lucrative career options. However, dropout rates at universities across the STEM fields remain high and many students do not finish their studies (Heublein et al., 2022). We define dropout as the decision to leave university completely without a degree (Heublein et al., 2017), which is different from the decision to change major or the university. For some students, the decision might be a wise correction in their personal career path. However, for most students, especially dropout has negative consequences concerning salary and well-being (Berlingieri et al., 2021).

The German university system places a significant emphasis on individual responsibility, with few mandatory elements and in some courses only one major exam per semester. Consequently, self-regulation and self-motivation

play pivotal roles, especially in STEM fields where the challenging nature of mathematics has been highlighted (Sithole et al., 2017). Performance problems in particular play a decisive role in the dropout decision (Heublein et al., 2017). However, dropouts often exhibit passive study habits and underutilize available university support (ibid.), preventing these students from truly integrating into the academic environment.

Both dropout and changing major must be seen as a multi-dimensional process with many different factors (ibid.) and the problem must be tackled from many different angles. Mindset research can contribute to this issue by starting at students' personal beliefs and personal experiences in the transition from school to university. Previous studies have observed a drop in motivation directly in the first semester (Dresel et al., 2013, Diederich et al., 2022), highlighting the necessity for targeted interventions and support structures in the first semester to enhance student success.

Theoretical foundation

What do we mean by "mindset" and how could it influence study success?

Mindset according to Dweck describes the belief (also called implicit or lay theory) that personal characteristics, such as intellectual abilities, are either largely predetermined and unchangeable (fixed mindset or entity theory) or can be developed through effort, good strategies, and the support of others (growth mindset or incremental theory) (Dweck & Yeager, 2019). Years of research have shown that a fixed mindset is accompanied by unfavorable learning behaviors and attitudes and, under some conditions, poorer academic performance, especially in situations with challenges and setbacks like the first semester of university (for an overview see e.g. ibid.; Dweck 1999; Burnette et al 2013). Dweck & Yeager (2019) theorize that the main mechanisms lie in a negative connotation of effort, performance avoidance goals and helpless responses (attributions) after setbacks (the "Mindset Meaning System"). We could replicate these correlations for our context at the start of STEM university studies (Diederich & Spatz, 2024). Other authors have proposed different path diagrams including e.g. stress, life satisfaction or proactive coping (see Yan & Schuetze, 2023 for an overview).

Mindset is typically measured by three to eight items about the malleability of intelligence (e.g. "You can learn new things, but you can't really change your basic intelligence", Dweck, 1999). There has been some debate about the validity of this "intelligence mindset" scale (e.g. Limeri et al., 2020a, Yan & Schuetze, 2023), as the word intelligence in the items is understood differently by students and they score differently depending on their definition of intelligence (Diederich & Spatz 2024). Several research teams developed subject specific scales (e.g.

physics (Malespina et al., 2022), STEM (Limeri et al., 2023), language (Lou and Noels, 2017)) to better describe the relevant beliefs of students in the specific context. Our research group developed a scale specifically for students studying a STEM subject at university, which we validated in a large sample with the "Mindset Meaning System" (Diederich & Spatz, 2024). In general, despite the debate about how to measure it best, the evidence for the role of mindset in motivation is strong.

The role of mindset in the decision to change major or drop out

The decision to change major or dropout usually takes place in the first semesters at the start of university (Heublein et al., 2017). To describe the decision process more closely, Bäulke et al. (2022) developed a five-stage model for the decision to drop out or to change major: Starting with a sense of non-fit perception (the first phase), thoughts become increasingly concrete (2. thoughts of quitting/changing, 3. deliberation, 4. information search) up to the final decision (5.). Especially the first phase both for drop out and changing mayor correlates highly with several emotional, motivational, and behavioral variables (ibid.). Therefore, mindset could play a role in this early process. In our previous studies (Diederich et al., 2022), we found subject specific fixed mindset to predict this non-fit perception for intentions to change mayor. For intentions to drop out, the correlation was small and not significant. In general, intentions to drop out have shown to be a good and easily accessible predictor of actual drop out (e.g. Fleischer et al., 2019, Dewberry & Jackson, 2018).

Mindset and subject specific academic self-concept

Academic self-concept is an important predictor for success at university (Robbins et al., 2004, Zajacova et al., 2005) and specifically for intentions to change major or drop out (Bäulke et al., 2022, Dresel et al., 2013). It is closely related to mindset, as both can be located on the expectancy side of personal beliefs and tendencies (see Dresel & Lämmle, 2017 for expectancy-value model of motivation including mindset (implicit theories)). While the academic self-concept represents a – not necessarily realistic – belief about the height of your personal academic abilities, mindset is about the possibility to change these abilities fundamentally (ibid.). In a previous study, we found both constructs to be uncorrelated at the start of the first semester (N=805, Diederich et al., 2022). According to Dweck (1999), we would expect this to be the case in a situation without a direct threat. Only in a case of setbacks, stronger higher fixed beliefs should lead to a decrease in confidence (a lower academic self-concept), as the setback might be attributed internally e.g. on a lack of talent (ibid.). Although we reported a decrease of academic self-concept over the first two month at university ($N = 201$, $d = -.16$), this decrease was not correlated with the mindset at the start of the semester, which does not the support the hypotheses that a growth mindset would protect academic self-concept. However, we also reported an increase

in fixed beliefs over this period. At the second measurement, higher subject specific fixed beliefs correlated with a lower academic self-concept (r = -.24). This could mean that mindset itself is influenced through academic challenges (as discussed by Limeri et al., 2020b) and only a robust growth mindset could protect academic self-concept.

How can we change the mindset of students?

Research indicates the malleability of mindset towards a conducive growth mindset, with early studies involving complex interventions (e.g., Blackwell et al., 2007) paving the way for more scalable approaches. Short online interventions, exemplified by the U.S. National Study of Learning Mindsets in the 9th grade (Yeager et al., 2019, benefits on GPA only for subgroup of 6320 low-achieving students, higher enrolment in advanced math courses especially for high-achieving students) and the U-Say experiment in Norway (Rege et al., 2021, benefits on challenge-seeking behavior), have emerged.

Controversies surround the effectiveness of such interventions in influencing academic achievement. Recent meta-analyses published in Psychological Bulletin in 2022 present conflicting conclusions on the robustness of evidence in mindset intervention research. The research group around Sisk, MacNamara and Burgoyne (Sisk et al., 2018; Macnamara & Burgoyne 2023a) heavily criticize the value of such programs („… positive results are rare and possibly spurious due to inadequately designed interventions, reporting flaws, and bias.“, ibid. S. 1). In contrast, Burnette et al. (2023) offer a more nuanced, positive assessment, acknowledging varying effectiveness but noting positive effects on academic outcomes, mental health, and social functioning, particularly when interventions are tailored to those expected to benefit the most.

Tipton et al. (2023) critique the methodology of the Macnamara & Burgoyne (2023a), emphasizing the importance of considering heterogeneity within each study (see Macnamara & Burgoyne, 2023b for reply). For instance, the National Study of Learning Mindsets identifies larger effects for students in supportive environments, with teachers' mindset moderating treatment effects (Yeager et al., 2019, 2022a). Nevertheless, the discussion underscores the importance of maintaining realistic expectations regarding the benefits of these interventions. Central to the controversy is the definition of a mindset intervention. Macnamara & Burgoyne (2023) advocate for a focus on the growth mindset message. However, effective interventions extend beyond fostering the belief in the malleability of intellectual abilities; they also provide actionable strategies (e.g., Yeager et al., 2016). E.g., interventions combining a growth mindset message with stress research has shown promising results (Yeager et al., 2022b). Yan & Schuetze (2023) posit that combining strategy-focused interventions with growth mindset messages represents a progressive step. Our intervention aligns with this comprehensive approach.

Our Intervention

Through multiple iterations, we developed an online single-session intervention (60 to 90 minutes) specifically for students at the transition from school to a STEM university major (initially with physics majors, see Rehberg et al., 2023). First, we start with the message that studying at university is quite different from studying at school and taking control over your studies is essential. We build on the "brain is like a muscle"-metaphor used in previous studies (e.g. Yeager et al., 2016) and argue for this view by introducing students to studies about neuroplasticity. We describe learning as a biological growth process between neurons in the brain. Next, we introduce three principles from learning psychology (retrieval practice, interleaving and spaced repetition), which Dunlosky et al. (2013) summarize in their literature review as having the highest utility. We introduce students to these findings and argue that studies find these principles to be efficient exactly because they are challenging the brain. Difficulties and struggle are therefore presented as a necessary part of learning and not as a sign of inadequacy. The intervention consists of four learning videos combined with interactive exercises that directly apply the principles it teaches. It includes stories of students about their experiences with difficulties containing growth mindset messages and a "saying is believing"-exercise (writing a letter to a struggling student) as in previous interventions (Yeager et al., 2016). The intervention is freely available (https://lernbar.uni-frankfurt.de/MaDi/BesserStud).

Research Question and Method

Our key research question is whether the online intervention can support students in their first semester by increasing their academic self-concept and decreasing their fixed beliefs as well as their intentions to change majors or drop out. To answer this research question, we conducted an online survey in introductory STEM courses at a German university in the 2022 winter semester (local ethics committee approved): during the first week of the semester (T1) and two months into the semester (T2). The sample consists of a total of 24 different majors with the largest groups consisting of approx. 10% of the sample (see table 1 for details). Students participated voluntarily during lecture time. At T1, the sample of students who properly completed the survey was n_1=1505 (approx. 80% of students present). At T2, the sample decreased to n_2=538 students as attendance decreased as well. 396 students from sample T1 could be identified in sample T2.

Directly after T1, students had the opportunity to voluntarily participate in the intervention, promoted as a study techniques course. Of the 395 in the matched group, 88 students completed the intervention fully (Intervention Group, IG) and 279 students had no contact with the material (Baseline, BL, self-reported). While the matched group is a self-selection compared to the sample at T1 (higher

percentage of women and native German speakers, better entrance grades), there is no significant difference between IG and BL.

Table 1. Summary of Sociodemographic data for different groups.

	T1	Matched T1 and T2		Matched BL	IG struggle subgroup	BL struggle subgroup
N	1505	396	88	279	21	59
Ø entrance grades (4.0 to 1.0 (best))	2.13	1.88***	1.88	1.86	2.19	2.07
% female	29%	34% (n.s.)	33%	36%	47%	51%
% first generation	33%	28% (n.s.)	32%	27%	38%	37%
% no parent speaks german	34%	24%***	24%	22%	14%	32%
Most frequent majors (N)	CS (185), ME (172), CI, Bi (97), EE (95)	ME (42), Ph (34), CS (32), EE, MT (29)	ME, MT (9), IE (8), EE (7), EE, Sy, MT, Ph (5)	ME (29), Ph (26), EE, CS (22), CI (21)	Med (4) Sy, Ph (3) Bi, BE (2)	ME (7), Bi, Ph (6), Ma, MT, TD (5)

BE: Biomol. Engineering, Bi: Biology, CI: Civil Engineering, EE: Electrical Engineering, IE: Industrial Engineering, CS: Computer Science, Ma: Mathematics, ME: Mech. Engineering, MT: Mechatronics, Ph: Physics, Sy: Inform. systems technology, TD: Teaching degree (STEM) *** < .1%, u-test (entrance grade) or chi-square-tests comparing T1 and the matched group.

Additionally, we asked students at T2 whether they had difficulties with the semester study content in the first weeks. 80 students self-reported continuous large difficulties (the struggle subgroup), of which 21 participated in the intervention (Measured with one item: "Did you have problems with the content of your studies in the first weeks?" and three possible answers: "Yes, but it is getting better.", "Yes, and this stays that way.", "No", self-developed oriented at Limeri et al., 2020b). N was underpowered for this subgroup (e.g.:). The subgroup is again a selection compared to the matched group (higher percentage of women and native German speakers, worse entrance grades)

Measures

For the survey we included two fixed mindset scales: the domain general "theories of intelligence" scale (Dweck, 1999, e.g. "You have a certain amount of intelligence and really can't do much to change it", 3 items,) and our subject specific mindset scale (Diederich et al., 2024, e.g. "I believe that some people are good in my field of studies and others will never really learn it, no matter how much they try.", 5 items,). We opted for selecting only items formulated in the fixed mindset direction due to practical considerations (limited time of about

8 minutes during lectures), and as items formulated in the growth direction are in danger to be socially desirable (Hong et al. 1999), following the approach of other large studies (e.g. three fixed mindset items in Yeager et al., 2019, for larger discussion of mindset scales see Diederich et al., 2024). We also assessed subject specific academic self-concept (Dweck 1999, translated by Dresel et al., 2013, e.g. ”I feel pretty confident… −” vs. ”I’m not very confident about my intellectual abilities, which I need in my field of studies”, 3 items,) and intentions to change majors or drop out (non-fit perception, Bäulke et al., 2022, e.g. “At the moment I don’t feel suitable for my major / for studying.”, each 3 items, (intentions to change majors, T2) and .87 (intentions to drop out, T2)). All scales range from 1 to 6 on a Likert scale.

Results

We compared the change over time in the mean and individual levels in IG and BL for both the complete matched group and the struggle subgroup (table 2). In both mindset measures, the IG reported a decrease in fixed mindset, while the BL reported no change over the period of two months. In the subject specific academic self-concept, there was a development to a higher self-concept in the IG, while the BL reported a small decrease. Both groups did not start at the same level (3.89 vs. 4.09), although the differences at the start are not significant. Effects are stronger for the struggle subgroup. In this group, the BL shows a clear decrease in academic self-concept, while the IG reported no changes.

On the individual level, we observe increases or decreases in all scales for the majority of students (except intelligence fixed mindset for BL), so there is a substantial development over the first two month. Differences between IG and BL are again larger for the struggle subgroup, where 31% of BL students report an increase in subject specific fixed beliefs (vs. 14% in IG) and 51% report a decrease in academic self-concept (vs. 24%).

Table 2. Change in both mindset measures and the subject specific self-concept in intervention group (IG) and control group (BL). Effect size d calculated for change in mean.

	Group (N)	mean at T1 (SE)		Effect size d (p-value) of	% incr.	% equal	% decr.
Intelligence fixed mindset							
Matched	IG (88)	3.12 (.12)	-0.33	-0.33 (.01)	20%	42%	38%
	BL (279)	3.15 (.07)	-0.06		21%	52%	27%
Struggle (subgroup)	IG (21)	2.97 (.27)	-0.44	-0.45 (.06)	19%	38%	43%
	BL (59)	3.04 (.14)	-0.03		24%	44%	32%
Subject specific fixed mindset							
Matched	IG (88)	2.96 (.09)	-0.15	-0.28 (.01)	22%	34%	44%
	BL (279)	3.01 (.05)	0.02		29%	41%	29%
Struggle (subgroup)	IG (21)	2.92 (.20)	-0.26	-0.39 (.07)	14%	38%	48%
	BL (59)	3.09 (.11)	0.03		31%	32%	37%
Subject specific academic self-concept							
Matched	IG (88)	3.89 (.13)	0.15	0.26 (.02)	40%	42%	18%
	BL (279)	4.09 (.06)	-0.10		23%	47%	30%
Struggle (subgroup)	IG (21)	3.41 (.24)	-0.05	0.40 (.07)	38%	38%	24%
	BL (59)	3.40 (.16)	-0.42		17%	32%	51%

Effect size d = . p-value of one-sided t-tests between IG and BL. % increase / decrease = perc. students with substantial change or ; else % equal. Likert scales from 1 to 6. SE: Standard Error.

Regarding intentions to change majors or drop out at T2 (table 3), the overall intentions were low. Only a minority reported values above 3, corresponding to any non-fit perception at all. In the struggle subgroup, more students report a non-fit perception. Here, we can find lower intentions both to change major and drop out in the IG. However, these differences are only significant for the difference in intentions to drop out with an effect size of (*p*-value without alpha error correction).

Table 3. Comparison of intentions to change majors or drop out in

intervention group (IG) and control group (BL) at T2.

Intentions to…		…change majors			…drop out		
	Group (N)	Mean	d (p-value)	% >3	Mean	d (p-value)	% >3
Matched	IG (88)	2.32	-.04 (.22)	15%	1.93	-.14 (.17)	7%
	BL (279)	2.36		15%	2.03		14%
Struggle (subgroup)	IG (21)	2.83	-.28 (.13)	33%	2.13	-.42 (.04)	14%
	BL (59)	3.17		41%	2.61		32%

% >3 is the percentage of students with mean scale values higher than 3 who report some intentions to change majors or drop out. p-values for u-tests due to deviation from normal distribution.

Discussion and conclusions

The online intervention successfully encourages the decrease of fixed beliefs and therefore the adaptation of a growth mindset, with small to medium effect sizes. Effects are stronger for students who experience ongoing difficulties (the struggle subgroup). Concerning subject specific academic self-concept, the development is also in the expected direction with an increase or at least no decrease (struggle subgroup) in academic specific self-concept in the IG. However, as the IG started on a lower level as the BL, both groups end up at a similar level and the evidence is not very strong here. Differences between intentions to change majors or to drop out are only found in the struggle subgroup, where the intervention group reported lower intentions.

In general, all results support the theorized direction. However, there are several important limitations. First, as students participated voluntarily, we might have a selection bias between IG and BL not shown in our sociodemographic variables. A larger study with an active control group design is needed to draw definitive conclusions. Second, as participants of the matched group are self-selected, results are not generalizable on the complete population. Nonetheless, as effects are stronger in the struggle subgroup with lower grades, the intervention might be more effective for the general student population. Third, due to small sample sizes, we could not differentiate between different majors, even though demands

and the degree of abstractness vary widely between them. Forth, all results are self-reports, and it is unclear, whether the intervention influenced actual student behavior. Nevertheless, results of this study indicate that a single online intervention can have a positive impact on students' mindset and subject specific academic self-concept in the early weeks of university. Due to the study design, results should be interpreted as exploratory.

As a Growth Mindset message fits well into the western individual culture and "just feels right" (Oysterman 2023), researchers should be especially careful in examining the effects of such interventions. Future research should follow best practice as closely as possible (a priori power analysis, active control group, pre-registration with definition of relevant subgroups, double blind experiments, random assignment, see Macnamara & Burgoyne, 2022a, p. 27) to avoid bias and measure effects on student mindsets, motivation and behavior additionally to outcome variables such as grades or dropout to clarify the mechanism through which the intervention is or is not working (see Yan & Schuetze, 2023 for possible future directions of mindset intervention research).

Considering the effect sizes, we cannot expect a large impact on student dropout through the intervention alone. After the intervention, there is still a large part of students with fixed beliefs (approx. 25% of students in IG have scale values > 3.5 for both mindset scales at T2). In interviews, we found especially the subject specific fixed beliefs to be persistent (Diederich et al., 2023). It is very difficult for students to develop a consisted growth mindset, especially in a challenging environment. Dweck herself describes it as a journey, as different situations can trigger fixed beliefs (Dweck, 2017). Therefore, we might expect larger results if the message and contents from the course are supported and repeated continuously during the semester. A study with computer science students usings tutors has demonstrated promising results with this continuous approach (Cutts et al., 2010). As voluntary participation rate is low (about 10%), students might benefit from a given time window to complete the course. As the time before exams is especially stressful, future studies should take a closer look on this.

Acknowledgement

The GROMM (Growing MINT Mindsets) project is funded by the German Federal Ministry of Education and Research. The authors are responsible for the content of this publication.

References

Bäulke, L., Grunschel, C., & Dresel, M. (2022). Student dropout at university: a phase-orientated view on quitting studies and changing majors. *European Journal of Psychology of Education, 37,* 853-876. https://doi.org/10.1007/s10212-021-00557-x

Berlingieri, F., Heigle, J., Pfeiffer, F., & Stichnoth, H. (2021). Analysen zu Kosten und Erträgen von Fachwechsel und Studienabbruch. In M. Neugebauer, H.-D. Daniel, & A. Wolter (Eds.), *Studienerfolg und Studienabbruch* (pp. 261 - 280). Springer VS.

Blackwell, L.S., Trzesniewski, K.H., & Dweck, C.S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: a longitudinal study and an intervention. *Child development 78 (1),* 246–263. https://doi.org/10.1111/j.1467-8624.2007.00995.x

Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M., & Finkel, E. J. (2013). Mind-sets matter: A meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin, 139(3),* 655–701. https://doi.org/10.1037/a0029531

Burnette, J. L., Billingsley, J., Banks, G. C., Knouse, L. E., Hoyt, C. L., Pollack, J. M., & Simon, S. (2023). A systematic review and meta-analysis of growth mindset interventions: For whom, how, and why might such interventions work? *Psychological Bulletin, 149(3-4),* 174–205. https://doi.org/10.1037/bul0000368

Cutts, Q., Cutts, E., Draper, S., O'Donnell, P., & Saffrey, P. (2010). Manipulating Mindset to Positively Influence Introductory Programming Performance. In *SIGCSE '10, Proceedings of the 41st ACM Technical Symposium on Computer Science Education* (pp. 431–435). Association for Computing Machinery. https://doi.org/10.1145/1734263.1734409

Dewberry, C., & Jackson, D. J. (2018). An application of the theory of planned behavior to student retention. *Journal of Vocational Behavior, 107,* 100–110. https://doi.org/10.1016/j.jvb.2018.03.005

Diederich, M., & Spatz, V. (2022). Welche Rolle spielt das Mindset beim Studienabbruch im MINT-Studium? Eine fächerübergreifende Erhebung im ersten Semester. *PhyDid B - Didaktik Der Physik - Beiträge Zur DPG-Frühjahrstagung, 1.* https://ojs.dpg-physik.de/index.php/phydid-b/article/view/1283

Diederich, M., Spatz, V., & Bauer, A. (2023). Auswirkungen einer Online-Intervention (Mindset, Lerntechniken) auf den Studieneinstieg. *Didaktik Der Physik - Beiträge Zur DPG-Frühjahrstagung - Jahresband 2023, 59–66.* https://ojs.dpg-physik.de/index.php/phydid-b/article/view/1333

Diederich, M., & Spatz, V. (2024). The role of domain-specific growth mindset (implicit theories) at the beginning of STEM university studies: Comparing new and old measures. *Preprint.* In *Effective Learning in Physics: From Contemporary Physics to Remote Settings* (Challenges in Physics Education series). Springer. https://doi.org/10.13140/RG.2.2.23851.82728

Dresel, M., & Grassinger, R. (2013). Changes in Achievement Motivation among University Freshmen. *Journal of Education and Training Studies, 1(2).* https://

doi.org/10.11114/jets.v1i2.147

Dresel, M., & Lämmle, L. (2017). Motivation. In T. Götz (Ed.), *Emotion, Motivation und selbstreguliertes Lernen (pp. 79–142).* Ferdinand Schöningh.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology. *Psychological Science in the Public Interest : A Journal of the American Psychological Society, 14(1)*, 4–58. https://doi.org/10.1177/1529100612453266

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development. Psychology Press*. https://doi.org/10.4324/9781315783048

Dweck, C. S., Yeager, D. S. (2019): Mindsets: A View From Two Eras. *Perspectives on psychological science : a journal of the Association for Psychological Scienc*e 14 (3), 481–496. https://doi.org/10.1177/1745691618804166

Fleischer, J., Leutner, D., Brand, M., Fischer, H., Lang, M., Schmiemann, P., & Sumfleth, E. (2019). Vorhersage des Studienabbruchs in naturwissenschaftlich-technischen Studiengängen. *Zeitschrift Für Erziehungswissenschaft, 22(5),* 1077–1097. https://doi.org/10.1007/s11618-019-00909-w

Heublein, U., Ebert, J., Hutzsch, C., Isleib, S., König, R., Richter, J., & Woisch, A. (2017). Zwischen Studienerwartungen und Studienwirklichkeit: Ursachen des Studienabbruchs, beruflicher Verbleib der Studienabbrecherinnen und Studienabbrecher und Entwicklung der Studienabbruchquote an deutschen Hochschulen. *Forum Hochschule: 2017, 1.* DZHW Deutsches Zentrum für Hochschul- und Wissenschaftsforschung.

Heublein, U., Hutzsch, C., & Schmelzer, R. (2022). Die Entwicklung der Studienabbruchquoten in Deutschland. (DZHW Brief 05|2022). *Deutsches Zentrum Für Hochschul- Und Wissenschaftsforschung GmbH (DZHW)*. https://doi.org/10.34878/2022.05.dzhw_brief

Hong, Y., Chiu, C., Dweck, C. S., Lin, D. M.-S., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology, 77(3),* 588–599. https://doi.org/10.1037/0022-3514.77.3.588

Limeri, L. B., Choe, J., Harper, H. G., Martin, H. R., Benton, A., & Dolan, E. L. (2020a). Knowledge or Abilities? How Undergraduates Define Intelligence. *CBE Life Sciences Education, 19(1).* https://doi.org/10.1187/cbe.19-09-0169

Limeri, L. B., Carter, N. T., Choe, J., Harper, H. G., Martin, H. R., Benton, A., & Dolan, E. L. (2020b). Growing a growth mindset: characterizing how and why undergraduate students' mindsets change. *International Journal of STEM Education, 7(1).* https://doi.org/10.1186/s40594-020-00227-2

Limeri, L. B., Carter, N. T., Lyra, F., Martin, J., Mastronardo, H., Patel, J., & Dolan, E. L. (2023). Undergraduate Lay Theories of Abilities: Mindset, universality, and brilliance beliefs uniquely predict undergraduate educational outcomes. *CBE—Life Sciences Education, 22(4).* https://doi.org/10.1187/cbe.22-12-0250

Lou, N. M., & Noels, K. A. (2017). Measuring Language Mindsets and Modeling Their

Relations With Goal Orientations and Emotional and Behavioral Responses in Failure Situations. *The Modern Language Journal, 101(1), 214–243.* https://doi.org/10.1111/modl.12380

Macnamara, B. N., & Burgoyne, A. P. (2023a). Do growth mindset interventions impact students' academic achievement? A systematic review and meta-analysis with recommendations for best practices. *Psychological Bulletin, 149(3-4)*, 133-173. https://doi.org/10.1037/bul0000352

Macnamara, B. N., & Burgoyne, A. P. (2023b). A spotlight on bias in the growth mindset intervention literature: A reply to commentaries that contextualize the discussion (Oyserman, 2023; Yan & Schuetze, 2023) and illustrate the conclusion (Tipton et al., 2023). *Psychological Bulletin, 149(3-4),* 242–258. https://doi.org/10.1037/bul0000394

Malespina, A., Schunn, C. D., & Singh, C. (2022). Whose ability and growth matter? Gender, mindset and performance in physics. *International Journal of STEM Education, 9(1).* https://doi.org/10.1186/s40594-022-00342-2

Oyserman, D. (2023). Culturally fluent theories, metascience, and scientific progress: A case example: Commentary on Macnamara and Burgoyne (2023) and Burnette et al. (2023). *Psychological Bulletin, 149(3-4),* 220–228. https://doi.org/10.1037/bul0000376

Rege, M., Hanselman, P., Solli, I. F., Dweck, C. S., Ludvigsen, S., Bettinger, E., Muller, C., Walton, G., Duckworth, A., & Yeager, D. S. (2021). How Can We Inspire Nations of Learners? An Investigation of Growth Mindset and Challenge-Seeking in Two Countries. *American Psychologist, 76(5),* 755–767. http://dx.doi.org/10.1037/amp0000647

Rehberg, J., Wilhelm, T., Spatz, V., & Diederich, M. (2023). Längsschnittstudie und Online-Workshop zum physikspezifischen Mindset. V. Vorst, H. (Hrsg.). *Lernen, Lehren und Forschen in einer digital geprägten Welt, Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Aachen 2022.* https://gdcp-ev.de/wp-content/uploads/securepdfs/2023/07/E25_Rehberg.pdf

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130(2),* 261–288. https://doi.org/10.1037/0033-2909.130.2.261

Sisk, V. F., Burgoyne, A. P., Sun, J., Butler, J. L., & Macnamara, B. N. (2018). To What Extent and Under Which Circumstances Are Growth Mind-Sets Important to Academic Achievement? Two Meta-Analyses. *Psychological Science, 29(4),* 549–571. https://doi.org/10.1177/0956797617739704

Sithole, A., Chiyaka, E. T., McCarthy, P., Mupinga, D. M., Bucklein, B. K., & Kibirige, J. (2016). Student Attraction, Persistence and Retention in STEM Programs: Successes and Continuing Challenges. *Higher Education Studies, 7(1),* 46. https://doi.org/10.5539/hes.v7n1p46

Tipton, E., Bryan, C., Murray, J., McDaniel, M., Schneider, B., & Yeager, D. S. (2023). Why Meta-Analyses of Growth Mindset and Other Interventions Should Follow Best Practices for Examining Heterogeneity: Commentary on Macnamara and

Burgoyne (2023) and Burnette et al. (2023). *Psychological Bulletin, 149(3-4),* 229–241. https://doi.org/10.1037/bul00003846

Yan, V. X., & Schuetze, B. A. (2023). What is meant by "growth mindset"? Current theory, measurement practices, and empirical results leave much open to interpretation: Commentary on Macnamara and Burgoyne (2023) and Burnette et al. (2023). *Psychological Bulletin, 149(3-4),* 206–219. https://doi.org/10.1037/bul0000370

Yeager, D. S., Romero, C., Paunesku, D., et al. (2016). Using Design Thinking to Improve Psychological Interventions: The Case of the Growth Mindset During the Transition to High School. *Journal of Educational Psychology, 108(3),* 374–391. https://doi.org/10.1037/edu0000098

Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., Tipton, E., Schneider, B., Hulleman, C. S., Hinojosa, C. P., Paunesku, D., Romero, C., Flint, K., Roberts, A., Trott, J., Iachan, R., Buontempo, J., Yang, S. M., Carvalho, C. M., . . . Dweck, C. S. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature, 573(7774),* 364–369. https://doi.org/10.1038/s41586-019-1466-y

Yeager, D. S.; Dweck, C. S. (2020): What can be learned from growth mindset controversies? *The American psychologist 75* (9), S. 1269–1284. https://doi.org/10.1037/amp0000794.

Yeager, D. S., Carroll, J. M., Buontempo, J., Cimpian, A., Woody, S., Crosnoe, R., Muller, C., Murray, J., Mhatre, P., Kersting, N., Hulleman, C., Kudym, M., Murphy, M., Duckworth, A. L., Walton, G. M., & Dweck, C. S. (2022a). Teacher Mindsets Help Explain Where a Growth-Mindset Intervention Does and Doesn't Work. *Psychological Science, 33(1),* 18–32. https://doi.org/10.1177/09567976211028984

Yeager, D. S., Bryan, C. J., Gross, J. J., Murray, J. S., Krettek Cobb, D., H. Santos, P., Gravelding, H., Johnson, M., & Jamieson, J. P. (2022b). A synergistic mindsets intervention protects adolescents from stress. *Nature, 607(7919),* 512–520. https://doi.org/10.1038/s41586-022-04907-7

Zajacova, A., Lynch, S. M., & Espenshade, T. J. (2005). Self-Efficacy, Stress, and Academic Success in College. *Research in Higher Education, 46(6),* 677–706. https://doi.org/10.1007/s11162-004-4139-z

Enhancing STEM Teaching in Higher Education: A Study of Project-Based Pedagogical Development for Novice Teaching Assistants

Rikke Frøhlich Hougaard, Karen Louise Møller

Aarhus University, Denmark

Growing demands for STEM teaching in higher education to adopt evidence-based pedagogical practices highlight the importance of professional development for STEM teachers. Research on previous efforts has shown that it is not simple to support teachers in adopting new pedagogical evidence in own practice. This study reports on the development and implementation of a pedagogical course designed for Ph.D.-students serving as teaching assistants (TAs) in STEM higher education. The course was designed as project-based learning where TAs work with small-scale projects in own teaching . Data collection includes both quantitative and qualitative data from pre- and post-course questionnaires, thematic analysis of project reports, and semi-structured interviews with TAs. The paper aims at widening our insight into the challenges and desires for professional growth experienced by novice TAs, and investigate the learning outcomes and meaning making related to working with a small-scale development project in own teaching. The findings are discussed in relation to professional growth and sustainable adaptation of evidence-based pedagogical practices by novice TAs in STEM higher education.

Keywords: University teaching, professional development, teaching assistants

Introduction

There is growing demands for STEM teaching in higher education to prioritise student-centred and evidence-based pedagogical practices including inquiry-based teaching, active learning, and research-based teaching (Handlesman et al., 2004; Penprase, 2020; Wieman, 2014). Despite evidence that active learning and inquiry-based approaches significantly increase student outcomes, traditional instructional strategies such as cookbook-like laboratory-teaching and lecturing often prevail in STEM university classrooms (Nielsen & Hougaard, 2018; Nielsen et al., 2020; Freeman, 2014). This highlights the importance of professional development for STEM teachers, in higher education.

In this study we report on the development of a pedagogical course (3 ECTS) for Ph.D.-students who work as teaching assistants (TAs) within STEM disciplines

at a Danish university. Referring to models for teacher professional growth the course was designed as project-based learning with the individual TA's experimentation and observation of students learning outcomes in own teaching practice at the centre of the course design. During the course the TAs plan, implement, and report a small-scale development project in own teaching (25 hours workload).

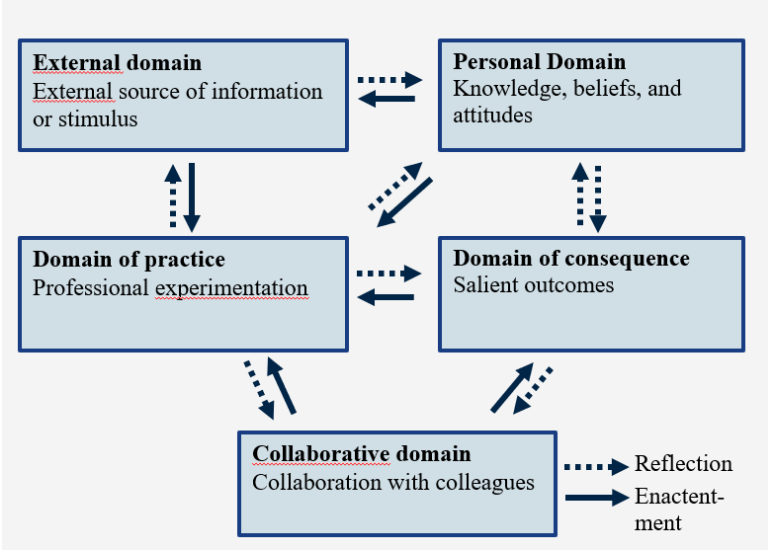
The present paper focuses on widening our insight into the challenges and desires for professional growth experienced by the TAs, which challenges, outcomes and instructional strategies they address in their projects, and how they perceive their learning outcomes after completing the project in their own teaching practice. This will inform the further development of the course and the alignment of pedagogical development efforts in STEM university teaching at different career levels.

Theoretical Background

Research on previous efforts has shown that it is not simple to support teachers in adopting new evidence based instructional strategies and ideas in own practice. Teachers and students who are inexperience with active learning strategies are likely to perceive the learning outcome higher from passive than active learning strategies (Deslauriers, 2019), and STEM teachers often rely more on personal experience than on empirical evidence in relation to teaching (Andrews & Lemons, 2015). Recent data suggest that STEM teachers who have been exposed to active learning as a student or worked in a team that used active learning are more likely to adopt such practices in own teaching (Apkarian, 2021).

The complexity in teacher professional growth is reflected in the "interconnected model of professional growth", according to which professional growth is best mediated when teachers are supported in experimenting in their own practice and experiencing desired outcomes in student behaviour and learning (Clarke & Hollingsworth, 2002). Referring to the concept of teacher agency (Biesta et al., 2015) the ability of a teacher to direct their professional growth and enhance their teaching practice is not an individual capacity but highly linked to their environment. Nielsen (2012) emphasise the importance of collaboration with colleagues to stimulate and support reflections and inquiries into own practice (Figure 1).

Figure 1. A model illustrating five domains which mediate professional growth via reflections and enactments. The model is based on Clarke and Hollingsworth (2002) and Nielsen (2012)



For Ph.D. students working as TAs in STEM university teaching it is not simple to establish an environment which stimulate professional development in relation to teaching, since their main occupation is research (Wright 2011). Initiatives such as the Learning Assistant model, which focus on local needs and context, has demonstrated to be highly effective in promoting professional growth and adoption of evidence-based instructional strategies for TAs in such environments (Odden et el., 2022; Top et al., 2018).

In this study we investigate how the planning, implementation, and reporting of small-scale projects in the context of a faculty specific pedagogical course may promote professional growth and adoption of evidence based instructional strategies for novice STEM TAs. This leads us to the following research questions.

1. Which challenges and desires for professional growth do TAs in STEM university teaching experience in their own teaching practices?
2. Which challenges, outcomes, and instructional strategies are addressed by TAs in the small-scale development projects?
3. Which learning outcomes and meaning-making can be identified after TAs complete a small-scale development-projects in their own teaching practice?

Methods

The course described in this study has been developed and adjusted in collaboration with participating TAs and faculty management. The study operates with an iterative Design Based Research approach (Barab & Squire, 2004) where data from course evaluations and assessments of TAs' work was used to inform the development of the course through several iteration. The data reported result from three iterations of the course in 2021-2023 including a total of 209 TAs.

For each iteration of the course, data was collected from a pre-course survey administrated to TAs online (n=209, week1-2), open reflections from online assignments (n=204, week 2), project reports (n=204, week 14-15), and a post-course survey (n=183, week 16) (week numbers refer to figure 2).

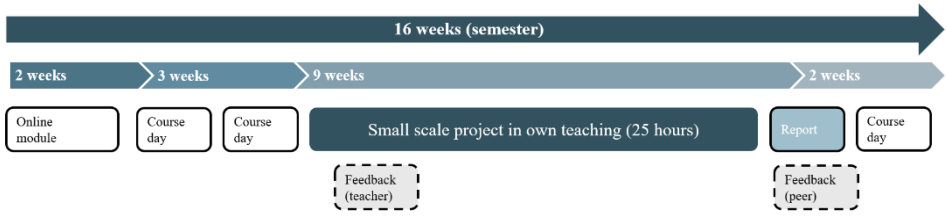
Quantitative as well as qualitative data related to RQ1 were collected from TAs' answers in the pre-course survey and in written reflections on experienced challenges in own teaching and desired outcomes for own professional growth. Data to answer RQ2 and RQ3 were obtained from project reports, from closed category and open reflections in a post-course questionnaire, and from semi-structured interviews with TAs from two iterations of the course (n=8) (Kvale and Brinkmann, 2009). Interviews were performed 1-8 months after the TAs completed the course, using a semi-structured interview guide including items related to perceived learning outcomes and opportunities and barriers to own professional growth as a TA. Informants represented 6 different departments (5 male, 3 female), and two different iterations of the course. They are referred to as participant 1-8, (P1-P8) in the following.

Challenges described in open reflections and addressed in the project reports were coded with the final categories, after an iterative process, and with a final quantitative coding with two coders calculating inter-rater reliability. Each response could be coded with several categories. Likert-scale items were analysed by frequency analysis and cross tabulations (Braun & Clarke, 2006).

Course Design

The course was designed as a 16-week blended-learning format including two initial weeks of online learning, three on-campus course days, and a project period (approximately 9 weeks) in between the on-campus course days 2 and 3 (Figure 2).

Figure 2. Schematic representation of the course design and timeline.



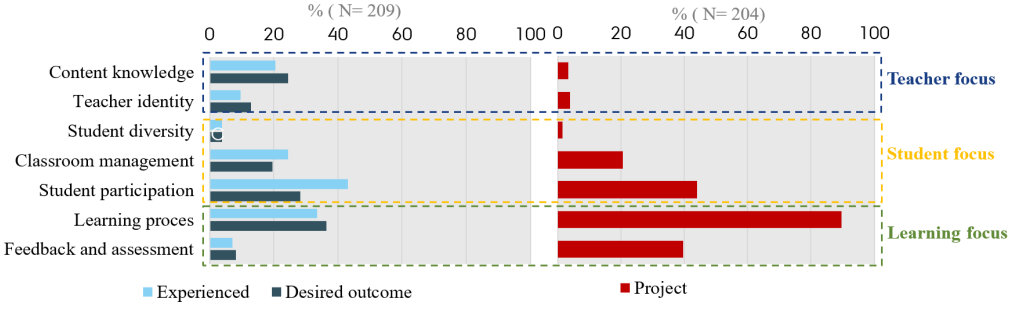
The course included mediation of theories and evidence-based teaching (external domain). Teaching and learning activities used in the course were designed to expose TAs to evidence-based teaching practices and to scaffold their reflections (personal domain). Templates, guidelines, and feedback from teachers were designed to guide the TAs' planning, implementation and reporting of the small-scale projects (domain of practice), and to support their identification of outcomes from their experimentation (domain of consequence). Group work and peer-feedback was used to ensure continuous collaboration among TAs (domain of collaboration).

Results

The TAs in the course were involved in various forms of STEM teaching and the projects addressed laboratory teaching (35%), theoretical exercises (42%), lectures (5 %), computer classes (15 %), project work (2 %) and field trips (1%).

The coding of challenges in own teaching and desires for professional growth reported by the TAs revealed seven categories, which could be divided into three main themes:1) Own performance and appearance (teacher focus), 2) Student characteristic and behaviour, (student focus) and 3) student understanding and learning (learning focus) (Figure 3).

Figure 3. Thematic analysis and quantification of challenges described by the TAs.



An example from teacher focus: “I find interacting with the students challenging. I would like to learn how to fully step into the role as a TA, so that the students feel they can come and ask me questions.”

An example from this student focus: “a group of students would not keep silent when other groups were presenting. It was very disrupting for everyone else.”

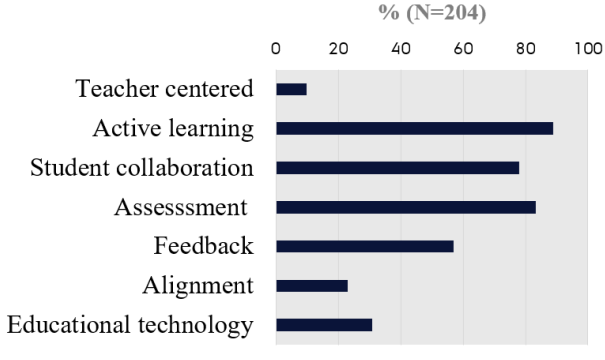
An example from learning focus: “I had doubt about whether the concepts were not fully understood by the students.”

In their open reflections in the beginning of the course, 20 % of the TAs mention challenges related to own content knowledge 40 % mention student participation and 30 % mention students’ learning (Figure 3, left panel, light blue bars). Their desired outcomes from participation in the course largely correlated with the challenges they experience (Figure 3, right panel, dark blue bars).

All categories identified in the open reflection were represented in the project reports submitted in the last week of the course. As evident from Figure 3 (right panel), the majority of projects (90 %) addressed challenges related to students’ learning process, compared to 33 % in the open reflections. 40 % of the TAs mention challenges related to feedback and assessment in the report, while only 7 % mention this in their initial reflections. Thus, TAs demonstrate an increased awareness of learning focused approaches after working with the small-scale projects.

The majority of the instructional strategies applied in the projects could be categorised as student-centred and 90 % of the projects adopted one or more of the evidence-based strategies covered during the course. More than 50 % of the projects included teaching formats which were demonstrated during the course days and only 10 % of the projects adopted solely “teacher centred” strategies, including e.g. performing a lecture (Figure 4, right panel).

Figure 4. Instructional strategies described in the project report.

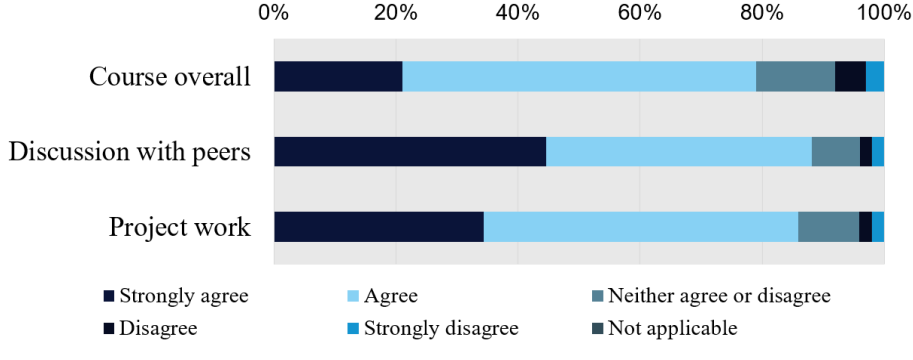


The quantitative data presented above may reflect a shift in the TAs' beliefs about teaching towards a more learning-focused perspective during the course, however, many other factors may have influenced their choice of focus and instructional strategies. To investigate this, eight TAs were interviewed 2-8 months after they completed the course.

In the interviews all interviewees mentioned that the course content and activities greatly inspired them, but that their choice of teaching strategy was based on own reflections on course alignment and current teaching format (eight interviewees) and how their own teaching could be improved (five interviewees). One of the interviewees explained that she was: "Highly inspired by what we did on the course days... it was easy to convert this to something the students found difficult" (P4). The templates, instructions, and guidance from the teachers were experienced by all interviewees to support their work with the small-scale projects, but it did not restrict their choice of challenge or instructional strategy. In particular, it was emphasised that this stimulated their individual reflections and their ability to inquire into own practice and try out something new. One TA says: "The template … made you think about whether the choices you make, make sense." (P3)

In the post-course questionnaire 89 % of the TAs reported that it was valuable to discuss their teaching with peers (figure 5). It was confirmed by all interviewees that these peer-discussions stimulated their reflections on own practice.

Figure 5, Perceived learning outcomes from the course.



Only three of the interviewees reported that discussions with peers influence their choices related to the projects whereas six felt that their teaching context and expectations from more senior colleagues restricted the extent of changes and experimentation in their projects: "The teaching is set in stone, and you are not expected to think about which challenges the students are facing." (P3)

When asked about their perceived learning outcome of the teaching and learning activities in the course, 86 % of the students found that the planning, implementation, and reporting of the small-scale project supported their learning

(agree and strongly agree, Figure 5), which is slightly higher than the perceived outcome of the course overall. All interviewees expressed that the most salient learning outcomes from working with the project was that they experienced being able to make intentional choices in own teaching based on reflections on students' learning, e.g. "Before I taught by intuition. Now I know something." (P1)

71% of the TAs reported a positive outcome of their project, demonstrating that they were able to influence the teaching format and make a difference for student learning. In the interviews several TAs expressed a belief that it is acceptable and important to experiment in their own practice also beyond the project, e.g.: "it is important for TAs and teachers in general to try out new things/approach in teaching. Not all things will work but it allows you to improve" (post-course survey). However, when the TAs were asked (1-8 months after the course) if they had or planned to adopt similar approaches in the teaching again, they all expressed that their willingness or ability to do so highly depended on the context of their teaching, e.g: "I could do it again. But again, it depends on the other teachers. For some, it can feel good and safe. Others I have only communicated with over email. I don't want to suggest anything new there." (P4)

Conclusions and Discussion

In the context of a pedagogical course in STEM university teaching, 204 TAs were exposed to evidence-based teaching methods and developed and implemented small-scale projects in their own teaching. The small-scale projects played a pivotal role in the course design, with the intention that the project work should stimulate reflections and enactments in own teaching practice. Referring to Clarke and Hollingsworth (2002) and Nielsen (2012) this may in turn lead to persisting changes in TAs personal beliefs and classroom practices when they experience positive effects in own teaching.

Challenges described by the TAs in their project reports reflect that the majority of TAs adopted an increased focus on students' learning during their participation in the course, compared to their initial experiences and reflections on their own teaching. This shift in focus is primarily attributed to new knowledge about evidence-based teaching and personal experience with evidence-based teaching activities gained during the online module and course days (external domain, Figure 1).

In the context of the small-scale projects 90 % of the TAs adopted evidence-based student-centered instructional strategies and the majority experienced a positive outcome from experimenting in their own teaching practice. The TAs choice of instructional strategy was intentional and based both on their personal beliefs of what would be beneficial (personal domain) and the amount of autonomy and the influence they experienced in their role as a TA in the specific teaching

context (domain of practice).

The collaboration with peers in the context of the course (domain of collaboration) stimulated the TAs' reflections and inquiries into own practice but did not enhance their willingness to adopt new practises. In contrast, traditions and expectations in the department were determinant for the extent of change in practice.

Thus, our data suggest that in the context of the small-scale project new input from the course (external domain) and interactions with peers (collaborative domain) stimulated reflections on own practice and mediated a shift in their personal beliefs about teaching (personal domain). This led to enhanced willingness to change their teaching practice (enactment). Factors restricting their willingness to change practice primarily related to the teaching culture at the workplace (domain of practice).

The TAs reported a high learning outcome from planning, implementing and reporting of the projects. They experience to be able to influence the teaching format and make a difference for students' learning. The most salient outcomes recognized by the TAs related to being able to identify challenges in own teaching, find solutions, and reflect on improvement in their practice with focus on student-centered learning. This suggests that in the context of the course the TAs adopted a new view of their own role and were able to make intentional choices in their teaching practice.

Based on the above we conclude that the project-based approach in this initiative to support professional growth for novice STEM TAs was successful in making TAs adopt ideas about evidence-based teaching in their personal beliefs and evidence-based teaching strategies in their own teaching practice. This transformation was mediated primarily via the TAs personal experience with evidence-based teaching as students in the course and from being supported in reflections and enactments in own teaching practices by peers and teachers in the course. These results are consistent with previous findings reported by Apkarian et al. (2021) and Clarke and Holingsworth (2002), respectively.

The TAs observed salient outcomes from experimenting in their own practice in STEM university teaching. Their persisting transformation to evidence-based teaching practices and development of teacher agency was, however, highly dependent on expectations from senior colleagues and departmental traditions where TAs often do not have much influence on the teaching (Biesta et al., 2015). In such environments, the learning assistant model, where professional development is linked to TAs participating in teaching teams in their own department, has been demonstrated to be highly effective in promoting adoption of evidence-based instructional strategies (Odden et el., 2022).

Our study underscores the importance of anchoring professional development efforts in the local teaching context and the current development of the course

focuses on establishing local collaboration between faculty staff and TAs during and after their participation in the course.

References

Andrews, T. C., & Lemons, P. P. (2015). It's personal: Biology instructors prioritize personal evidence over empirical evidence in teaching decisions. *CBE—Life Sciences Education,* 14(1), ar7.

Apkarian, N., Henderson, C., Stains, M., Raker, J., Johnson, E., & Dancy, M. (2021). What really impacts the use of active learning in undergraduate STEM education? Results from a national survey of chemistry, mathematics, and physics instructors. *PloS one*, 16(2), e0247544.

Barab, S., & Squire, K. (2016). Design-based research: Putting a stake in the ground. In Design-based Research (pp. 1-14). *Psychology Press.*

Biesta, G., Priestley, M., & Robinson, S. (2015). The role of beliefs in teacher agency. *Teachers and teaching*, 21(6), 624-640.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101.

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and teacher education,* 18(8), 947-967.

Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences*, 116(39), 19251-19257.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the national academy of sciences,* 111(23), 8410-8415.

Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R., ... & Wood, W. B. (2004). Scientific teaching. *Science*, 304(5670), 521-522.

Kvale, S., & Brinkmann, S. (2009). Interviews: Learning the craft of qualitative research interviewing. *sage.*

Nielsen, B. L. (2012). Science teachers' meaning-making when involved in a school-based professional development project. *Journal of science teacher education,* 23(6), 621-649.

Nielsen, B.L. & Hougaard, R.F. (2018). Scaffolding students' reflective dialogues in the chemistry lab: Challenging the cookbook. *In ESERA 2017 proceedings.*

Nielsen, B. L., Hougaard, R. F., & Krægpøth, M. (2020). Forskningslignende laboratorieaktiviteter for 1. års studerende på universitetet – muligheder og udfordringer. MONA - *Matematik- Og Naturfagsdidaktik*, 2020(1), 19

Penprase, B. E. (2020). STEM Education for the 21st Century. *Springer Nature*.

Odden, T. O. B., Lauvland, A., Bøe, M. V., & Henriksen, E. K. (2023). Implementing the learning assistant model in European higher education. *European Journal of*

Physics, 44(3), 035701.

Top, L. M., Schoonraad, S. A., & Otero, V. K. (2018). Development of pedagogical knowledge among learning assistants. I*nternational Journal of STEM Education*, 5, 1-18.

Wright, M. C., Bergom, I., & Brooks, M. (2011). The role of teaching assistants in student-centered learning: Benefits, costs, and negotiations. I*nnovative Higher Education,* 36, 331-342.

.

STEM Based Astronomy Activities with Ancient Minoan Analogue Computer

Apostolos Xenakis[1], Christos Xenakis[3], Sarantos Psycharis[2], Konstantinos Kalovrektis[1]

[1]University of Thessaly, Greece

[2]School of Pedagogical and Technological Education, Greece

[3]Astronomy and Space Society, Greece

The period of the Minoan Civilization, the first great European Civilization, lasted from 3000 BC. until 1000 BC approximately. During this long period, shipping, agriculture and rural economy, water supply and drainage in large Palace structures, metallurgy and metalwork, pottery and seal carving, the art and color aesthetics of wall paintings, ceramics painting, lithography, micro sculpture and micro plastic were developed. In addition to the above, very important milestone is the development of Minoan scripting language, that is the Phaistos hieroglyphs, Linear A, and B. However, the less known achievements, but also very important has to do with the numbering system, the use of fractions with the analogical symbols, as well as the design and use of the first Analogue Computer, a radial small dimensions disk, the size of a palm, whose matrix is found in the location of «Paleokastro», near Sitia city in Crete. The use of this computer is multifaceted, because it functioned as a sundial, as a navigation instrument for finding the latitude of a place and as a system for predicting eclipses of the Sun and Moon! In this paper, we give background information related to the functionality and use of the analogue computer, as a scientific instrument, and we propose appropriate STEM Astronomy activities, which are supported by it. To this end, we argue to what extent, ancient Greeks did apply STEM. Our work focuses to measure the impact of using Minoan computer to the learning curve of students, working with Astronomy problems.

Keywords: Astronomy, STEM, Paleokastro Crete, Analogue Computer

Introduction

During the period of the Minoan Civilization, the inaugural phase of a prominent European culture, spanning approximately from 3000 BC to 1000 BC, various facets of societal progress unfolded. Throughout this protracted era, advancements were witnessed in navigation, agriculture, rural economy, hydraulic engineering, and the construction of extensive palatial complexes. Concurrently, there was a flourishing of metallurgy, metalworking, ceramics, seal engraving, as well as

artistic endeavors, manifesting notably in the chromatic aesthetics of frescoes, ceramics, and lithographs.

In addition to the aforementioned achievements, a significant milestone was the development of Minoan script, encompassing the hieroglyphics of Phaistos, Linear A, and Linear B. However, lesser-known yet highly noteworthy accomplishments pertain to the numerical system, the utilization of fractions with proportional symbols, and the design and application of the first Analog Computer. This pioneering device, a radial disc of small dimensions resembling a palm-sized disk, with its matrix located in the vicinity of Palekastro, near the city of Sitia in Crete, marked a substantial advancement in ancient computational technology. In Figure 1, we depict the Matrix representation of Palaikastro Analogue Computer, as retrieved by the Ancient Museum in Heraklion, Crete.

The Minoan Computer stands as indisputable evidence of astronomical knowledge grounded in millennia of meticulous observations and the application of mathematics. It attests to the adept utilization of such knowledge for the creation of precise lunar-solar calendars and the capability to predict eclipses, demonstrating a notable accuracy in forecasting both solar and lunar events. The Minoan Computer stands as a remarkably significant and rare archaeological find from the early to mid-2nd millennium BCE, discovered serendipitously. Due to its lack of origin from systematic archaeological excavation, dating it poses challenges. Its existence underscores that ancient Greeks, as exemplified by this advanced astronomical device, developed sophisticated astronomy based on precise mathematics, astronomical knowledge, an understanding of causality, and recognition of the laws of nature. Through such advancements, we gain insights into, interpret, and in some instances, can predict certain astronomical phenomena, such as lunar phases, seasonal changes, and eclipses. Moreover, the mechanism hints at the presence of high-level technology in Crete, a notion echoed in the myths of Daedalus and Talos.

The utilization of this computer was multifaceted, serving as a solar clock, a navigation instrument for determining the geographical latitude of a location, and a predictive system for forecasting solar and lunar eclipses. The Minoan civilization flourished from approximately 2800 BC to 1100 BC, primarily on the island of Crete. In 1898, near Palekastro, east of Sitia in the Lassithi region of Crete, two slate matrices were discovered (in Fig. 1). On one of the matrices, the frontal aspect of the plate depicts a disk with two carved rays forming a cross, surrounded by serrated edges. At the center, a female figure with raised hands holding flowers is portrayed, while a circle with a cross rests above a crescent moon.

Figure 1. Matrix representation of Palaikastro Analogue Computer [retrieved from Ancient Museum of Heraklion at: https://www.heraklionmuseum.gr/en/]



The circle exhibits a crenellated base with horizontal stripes. This archaeological finding, originating from a region linked to the Minoan civilization, offers intriguing insights into the symbolic and artistic expressions of this ancient culture. Furthermore, on the front of the figure, two incised needles (stylized as compass needles) and a semi-circular incised blade resembling a compass dial are present.

The artifacts, was assessed by Sir. Arthur Evans and the director of Heraklion Museum and Archeologist, Stefanos Xanthoudidis. Subsequently, these findings have been on display at the Heraklion Museum in Crete. The estimated age of the artifact is between 1500 BC and 1400 BC. Later investigations by the Dutch scholar Jan Velsink resulted in dating the Palekastro matrix to the period of 1875-1850 BCE. Michael Tsikritzis undertook the comprehensive analysis of the functionality of the artifact in 2010.

Technical Background

In this section, we will give a necessary background information related to the functionalities and use of the Paleokastro ancient analogue computer. Following, we will provide a brief discussion related to the STEM epistemology and computational thinking.

Computer's Characteristics and Functionalities

The Minoan Computer consists of two disks, the large and the small, the latter having a diameter almost half that of the former. The primary functions are carried

out using the large disk, while the small one, for the time being, is deemed more symbolic. It could be argued that the small disk symbolizes the laws of nature, signifying the regularities that rhythmically repeat themselves, such as the 28-day phases of the visible Moon and the 18-year cycles of eclipses. Regularity, the recurrence of astronomical phenomena with periodicity, are elements that led Paleolithic humans to perceive the existence of natural laws, prompting Pythagoras to term the Universe as the Cosmos.

The dimensions of the artifact (i.e. matrix) of the plate are 8.5 cm x 8.5 cm. As depicted in Fig. 1, the matrix consists of a radial disk, surrounded by 25 projecting equilateral triangles, uniformly spaced around its periphery. Each triangle is circumscribed by 5 dots, also equidistant from each other, as depicted in Figure 2. Two of the triangles have 3 beads, while one has none. The total number of beads is 116.

Figure 2. A more detailed view of the matrix



Within the interior of the disk, there exists a smaller concentric circle with a diameter slightly smaller than two-thirds (2/3) of the diameter of the outer disk. This inner circle is composed of continuous equidistant notches. An inscribed cross within the circles (the symbol of the Sun) divides the matrix into four similar quadrants (see Figure 1). On one axis of the cross, there are 10 and 11 dots evenly spaced on the two radii, while at the bottom, there is a crescent symbolizing the Moon, along with 28 equidistant beads on the outer circle, indicating a lunar month. On the other perpendicular double diameter, 10 and 11 dots are discernible on one radius, while 7 and 8 beads are present on the other radius. The inner circle of the artifact, i.e., the computer, is divided into two hemicycles, each having 29 and 30 incisions. As revealed by the research, these incisions represent two lunar months of 29 and 30 days, commencing and

concluding with a full moon. The outer gear has 56 dots arranged in a circular pattern.

The computations performed by the Minoan Computer are focused on the precise determination of dates, primarily with the aim of predicting weather conditions for agricultural activities, determining suitable times for launching ships, and possibly for planning hunting expeditions. It can be utilized as a sundial and also for determining geographical latitude. Ancient mariners, and almost until recent times, navigated safely along the Mediterranean by using a fixed latitude slightly north of Crete, voyaging from Cyprus to Rhodes, Crete, Sicily, and up to the Iberian Peninsula, following the pillars of Hercules. Geographical latitude can be easily measured by observing the sun's height with a vertical rod placed at the center of a disk at noon. Naturally, the use of tables, as later became common, greatly aids the user of an astronomical instrument, whether it be the Minoan Computer, the Antikythera Mechanism, or any device utilizing a sundial for a multitude of purposes, especially if the user can also measure and account for time.

According to research the last 15 years, it has been found that Palaikastro Computer can do the following amazing things:

1. It works as a sundial clock (finder), on sunny days and tell us the time with an accuracy of 6 minutes! (one Minoan hour = 56 min of our time, see Fig. 3a)
2. It calculates the approximate latitude of the user's location and on the equinoxes (21/3 and 21/9), the accuracy was great! (see Fig. 3b)
3. It could accurately predict solar and lunar eclipses! (see Table 1 and Figure 4)
4. It could record the Sun's exposure within the Celestial sphere, that is, the successive positions of the Sun at the same time of day, but on different successive dates during a year! Based on the calculation of Anallema, Minoans could determine the seasons of the year.

Figure 3. Time (left figure) and Latitude (right figure) estimation

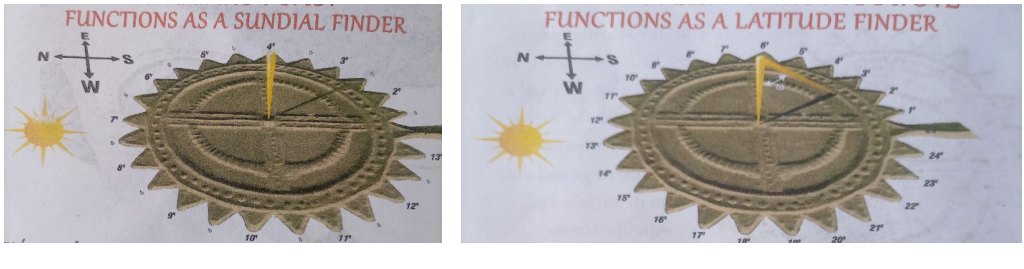
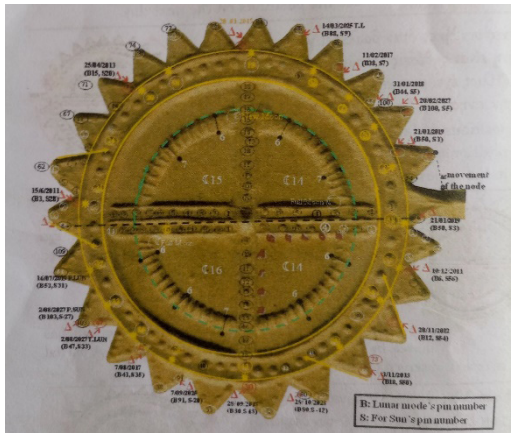


Table 1. Predicted, partial (P) or Total (T) Lunar or Sun Eclipses for the next 5 years!

Dates	Eclipse Type (Lunar or Solar)
28/10/23	Partial Lunar
18/9/24	Partial Lunar
14/3/25	Total Lunar
7/9/25	Total Lunar
28/8/26	Partial Lunar
2/8/27	Partial Solar
12/1/28	Partial Lunar
31/12/28	Total Lunar

Figure 4. Eclipse predictions



All the aforementioned functions were highly significant for the lives of the Minoans, contributing to the development of agriculture with knowledge of seasons, navigation with the ability for orientation and finding geographical latitude, fishing, hunting, forestry, or new tree planting, and the precise scheduling of daily tasks with accurate timekeeping. Another significant aspect of the uses of the Palaiokastro computer was its contribution to regulating the religious and worship practices of Minoan society.

In general, Greek civilization, dating back to the Minoan era, progressively advanced sciences and grounded its technological developments in scientific principles. As evidenced by the study of the specific computer, the Minoans developed advanced astronomy based on precise mathematics, careful observational astronomy, an understanding of periodicity and causality, and an acceptance of the existence of natural laws. Through these foundations, they comprehended natural phenomena and, at times, could predict periodic events such as seasons, phases of the moon, lunar and solar eclipses, and ultimately make informed speculations about seasonal climatic changes.

The existence of pioneering technology in Minoan civilization was well known through the myths of Daedalus and Talos. This is further substantiated by the Palaiokastro computer, which some researchers describe not as analog but as a digital device operating in the unary numeral system, where numbers are represented by repeating the same symbol as many times as the value of the number. In this system, the number 1 is represented by a single dot, line, or small cavity, and the number 5 is represented by five consecutive identical symbols, and so forth. These symbols, each representing a unit of time, can be considered as the bits of the prehistoric computer. As previously mentioned, these units are depicted on circles or straight segments (rays of the cross) because these shapes are the most suitable for the ancient user to manage periodic phenomena, such as hours of a day, months, years, phases of the Moon, eclipses of the Moon and Sun, etc.

STEM Framework and Engineering Design Process

There is a strong connection between STEM (Science, Technology, Engineering and Mathematics) and Computational Thinking (CT). The latter, is considered as a fundamental skill within the STEM disciplines. The problem-solving strategies and cognitive processes involved in CT are particularly relevant to STEM fields, where analytical thinking, algorithmic approaches, and logical reasoning play a crucial role. In STEM education, integrating CT within the learning process helps students develop, skills needed to solve complex problems, design systems and understand emerging patterns. CT provides a framework for approaching challenges systematically, breaking them down into manageable parts, and creating algorithms to devise solutions. These skills align well with the problem-solving nature of STEM subjects, making CT an integral part of STEM education and a valuable asset for individuals pursuing careers in STEM-related fields.

Computational Thinking (CT) entails dealing with real-world problems, developing and evaluating systems, and understanding human behavior by applying cognitive principles integral to computer science (CS). Widely acknowledged as a cross-disciplinary skill, CT fosters critical thinking in science, mathematics, and engineering, with a focus on system design. Despite being widely recognized, scholars contend that there is no exact definition for CT. Although CT is rooted in fundamental CS concepts, it includes supplementary practices such as abstraction, pattern recognition, simulation, problem decomposition, solution synthesis, verification, and prediction. The dimensions of CT are as follows:

- The ability to think algorithmically
- The ability to decompose problems
- The ability to generalize, to draw conclusions and find patterns among

data

- The ability to evaluate models
- The ability to think abstractly

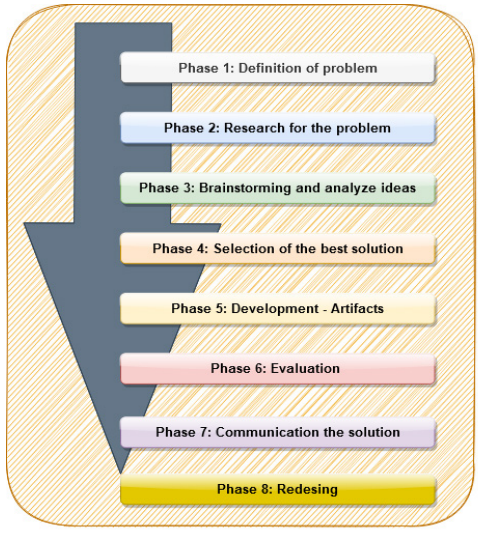
Computational Thinking involves the following seven practices:

1. Abstraction and Problem Decomposition (**ABS**)
2. Parallelism (**PAR**): Simultaneous processing of a task
3. Algorithmic Thinking (**ALG**): Discrete sequence of steps to complete a task
4. Decomposition (**DEC**): Breaking a problem into smaller and manageable parts, which can be solved independently and quickly.
5. Flow Control (**CON**): Related to algorithmic steps needed for completion
6. Testing and Debugging (**TAD**): Performing testing and fixing problems, while developing the system
7. Data Collection (**DAT**): Data collection, representation and analysis

In order to asses, in practice, CT, it is essential to formulate problems, preferably rooted in STEM disciplines. *Problem-Based Learning* (PBL) is an educational approach where students confront intricate problems, prompting them to navigate various cognitive challenges to arrive at solutions. In the context of Astronomy science course modules, and especially when students are required to design artifacts and use them for their experiments and observations, they are also required to collaborate in teams and progressively engage themselves with the dimensions of CT. Within the PBL structure, defining the problem is crucial, and equal emphasis is placed on the solution process. In PBL, an approach centered on problem-solving, students work collaboratively in groups, employing methods such as trial and error, model testing, hypothesis formulation, and debugging to derive solutions.

However, in STEM related courses, where artifacts and system design is highly involved, the Engineering Design Process (EDP) is also applied along the learning process. In essence, EDP, is a systematic and iterative approach, applied by engineers, to solve complex problems and reach innovative solutions. It is also a contemporary teaching method, appropriate for STEM use cases. In the majority of the cases, EDP may be combined with PBL in cases of interdisciplinary STEM projects. EDP involves a series of eight phases, to find a solution to a problem. It is an iterative process, with feedback loops. These phases usually include problem – solving processes, such as defining a problem, background research, specifying the requirements, teams brainstorming, evaluating solutions, developing a prototype and testing the solution.

Figure 5. The eight phases of EDP



In Fig. 5, we depict the eight EDP phases, which facilitate the process to reach a solution. In this work, students need to design and 3D print an artifact to simulate the majority of the functionalities of Palaikastro computer, to use it during a STEM astronomy scenario. The phases (P) are the following: **P1:** *Recognize the need or problem – Definition of the problem*, involving an understanding of the problem's nature and its connection to scientific and technological principles. **P2:** *Investigate the need or problem*, entailing online research (e.g., e-books, papers, databases, or AI tools) to gain a deeper understanding. Students explore existing solutions and seek new directions based on literature. **P3:** *Brainstorming and analyze ideas*, where students actively work on creating a prototype. **P4:** *Selection of the best solution*, involving the application of criteria to select the best option (or an acceptable one) from available alternatives. **P5:** *Develop a Prototype - Artifact*, requiring groups to finalize a prototype artifact. **P6:** *Test and evaluate the solution(s)*, with students conducting thorough testing and evaluations. **P7:** *Communicate the solution(s)*, encompassing the production of comprehensive test reports and the sharing of information about the prototype's functionality and usability. **P8:** *Redesign*, constituting a feedback loop where students may need to revise certain parts of their artifact or modify their solution to better align with optimization criteria.

In the following proposed STEM Astronomy didactic scenario, we will assign tasks to students, according to the EDP phases, and ask them to design and use their own Palaikastro artifact – computer, as produced by a 3D printer. To this end, we expect students to advance their knowledge in both pillars: The firsts related to the understanding and systemic design of the ancient computer and the second to use it and experiment with it, highlighting all of its functionalities.

Didactic Scenario and STEM Astronomy Activities

In this section, we discuss about the design and implementation of the STEM astronomy based didactic scenario, based on CT and EDP process. The following scenario and methodology applied in 45 students, 25 boys and 20 girls, with their ages varying between 16 and 18 years old, who are qualified for the finals of the annual Greek National Astronomy and Astrophysics competition, as organized by the Greek Astronomy and Space Society.

Scenario Description

The goals of the scenario are as follows:

1. Students should learn more about the amazing potentials of this ancient computer, as far as it's use is concerned as an instrument to measure time and latitude
2. Student should comprehend that this artifact had **at least three basic functions**:
    - The potential to perform direct measurements
    - The potential to perform predictions and immediate prediction verification (regarding Solar and Lunar eclipses)
    - The calculation of periodic celestial phenomena and therefore the verification of the existence of various natural regularities and laws.
3. Students should understand that in the early days of 18th and 19th century B.C., in a Mediterranean Island, in Crete, but also in the wider island complex of the Aegean Sea, a civilization was born that created science using the mathematics and was able to understand the laws - regularities of nature. This civilization used science and technology to construct artifacts that make their lives easier.

Following, we determine, according to the EDP phases, as depicted in Fig. 5, which are the goals per phase. In particular:

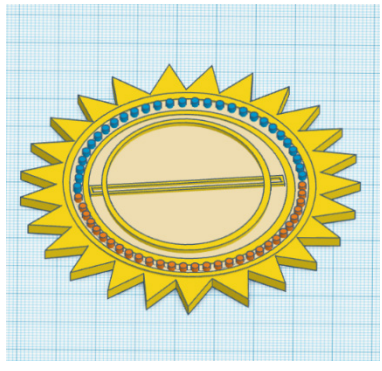
P1:

- Collection of high resolution photos of the find from the Archaeological Museum of Heraklion
- Enlarging them reveals many details about the finding
- With guidance, students can focus mainly on the triangles of the outer large circle and what they contain, as well as the two embossed needles and the dipper shaped forceps.

P2:

- Two copies (on thick cardboard) of the computer were made: one in physical size 8.5 cm X 8.5 cm and a larger one as supervisory material (demonstration computer).
- The ultimate goal is to reproduce the computer in a 3D printer for more tangible and realistic use by the students, as a hand-held computer in its natural size. (TinkerCAD 3D, see Fig. 6)

Figure 6. 3D Computer design in TinkerCAD (by Autodesk)



P3:

- Students were asked to count the triangles in the circumference, as well as the set of equidistant points contained within each triangle.
- It was found that 22 triangles had 5 equally distributed dots, 2 triangles had 3 dots and 1 triangle had no dots, confirming the findings of the expert scientists.

P4:

In the sequel, since the system could only function on sunny days as a sundial (sundial finder), and assuming that the maximum length of a day would not exceed 12.5 hours on average, it was requested to determine the accuracy, i.e. the sensitivity of the sundial (smallest unit of time the system could measure).

P5:

- An assumption was made to assign each triangle to 0.5 hours so that the total represents the 12.5 hours of maximum day length, if the time at sunrise and sunset is excluded.
- Also, the 5 dots of each triangle divide the 0.5 hours into smaller time units, each unit lasting approximately 6 minutes. This would also be the measurement accuracy of the sundial.

- According to the question, what was the duration of 1 hour back then, according to the Minoan sun watch, the answer was found to be about 56 minutes, as shown by the simple calculation: (116 dots) X (6min)= 696 Minoan min they correspond to 12.5 hours and each hour contains (696/12.5)~ 56 min, i.e. very close to our own time.

P6:

In the sequel, after the students got the basic knowledge about the operation of a horizontal sundial, the observational-experimental verification of the operation of the Minoan clock took place.

P7:

- In the center of the concentric circles, in the notch, one of the two needles (guide) was placed vertically and stabilized.
- Then one of the axes of the central cross was oriented in the North-South direction (see the picture), while the clock disc was kept horizontal.
- Then, the shadow of the needle (as a guidance) indicated the point on the radial disk that corresponded to the time of observation.
- The measurements were repeated and verified for several hours and at different times of the year. At the same time, the accuracy (sensitivity) of the instrument was checked every 6 min for every half hour of measurement.

P8:

- Carved into the matrix are three tools: two needles (as guidance) of a certain length and a semi-circular forceps (diabetes) that together with the computer can be used by a user to find the latitude of the place he is at.
- According to Fig. 3b, at the right hand side, students place vertically a pin hold the disc with the bearings, oriented in the North - South direction. Afterwards they record the angle procuded as the other side of the pin is tangent to the top of the vertical needle. This angle ω (Omega) records the latitude.

The following (3) scenario application conclusions are important:

1. Students test the artifact at different latitudes: Crete (35.3ºN), Athens (38ºN), Volos (39.4ºN), Thessaloniki (40.5ºN) and gave different but measurable results.
2. The measurements regarding **ω angle**, was done with a protractor, however, the Minoans should have measured it, with the help of perhaps some quarter circle (90 degrees), where they marked the angles of the

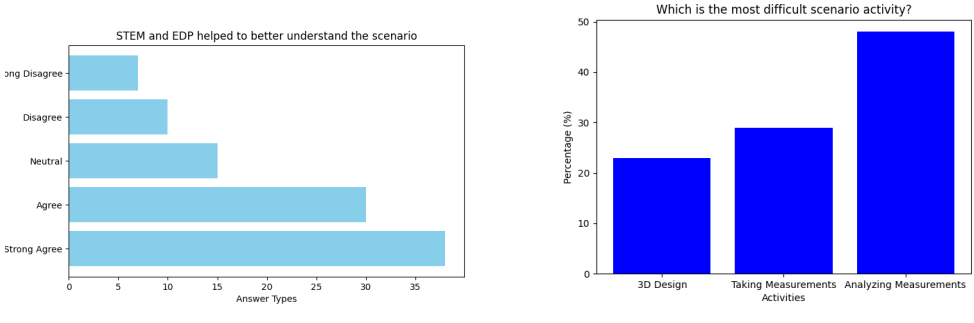
measured latitude. There is no information about that concept, other than the assumptions we can make.

3. If, for any reason, one moves north or south for a few days, then by the same process one could find the new latitude one is in, something very useful for merchants and sailors who traveled long distances, since with corresponding processes could return to the latitude where they started, moving in the opposite direction.

Results

We assess students' perspectives on STEM methodology, the Problem-Based Learning (PBL) process, and Engineering Design Process (EDP), in conjunction with the utilization of digital tools and smart learning environments. Employing a Likert 5-scale questionnaire, where responses range from 1 (strongly disagree) to 5 (strongly agree), our findings, as illustrated in Fig. 7a, indicate that a significant number of students express the view that engaging in project-based work and adhering to the EDP phases contributed to a deeper comprehension of how Palaikastro computer functions and how it is used in an experimental process related to Astronomy.

Figure 7. Students' answers related to STEM (a) and Activities (b)



In particular, the majority of the students strongly agrees that STEM and EDP scenario helped them to better understand, design and experimentally use the artifact, which simulates the functionality of the ancient Minoan analogue computer. However, they find it difficult to analyze measurements taken from the artifact.

Conclusions

In this work, our objective is to provide fundamental information regarding the basic functions and use of the ancient Minoan analog computer, as a scientific instrument. Additionally, our pedagogical approach is based on the contemporary method of EDP, wherein students design a *replica* of the Minoan computer, in 3D printer. Our basic research questions relates to which extend Ancient Greeks

applied STEM epistemology and computational thinking in their daily lives, addressing challenges related to their survival. Judging from the artifact's advanced, for the time being, functionalities, we argue that their minds work very scientifically.

Our intervention focuses on assessing the impact of using the Minoan computer, within the learning process of students, engaged in Astronomy problem solving modules. The students' involvement in the special learning project, gave them the opportunity to discuss why this particular technology was developed, what need demanded it and impose it, how it evolved over the centuries and what was the benefits of the Minoan society of a certain historical period. Due to the uniqueness of the finding, students also had the opportunity to discuss and understand the role that metal alloy or baked clay replicas of the utilitarian could have played in a large portion of social layers, such as merchants, sailors, military commanders etc. Following, students had the opportunity to familiarize themselves with a type of prototyping, 4000 years before our time. The possibility of replicating the computer from precious metals or cheap alloys or baked clay or pressed soft wood shavings mixed with resin allowed many professionals of the time to acquire and use it, having previously been trained in its functions.

By studying this particular artifact, students understand the usefulness of the computer's ability to make time and latitude measurements, predictions of natural phenomena over time, such as Moon and Sun's eclipses and to verify celestial periodic phenomena, according to physical laws and regularities. Finally, by participating in such a didactic-educational scenario, students were introduced to an interdisciplinary approach based on guided discovery and STEM methodology, *combining knowledge from history and archaeology*, *mathematics, astronomy, physics, IT and modern technology* (reproducing a copy of the artefact with a 3D printer).

References

Blomberg, M., & Henriksson, G. (1996). Minos Enneoros : Archaeoastronomical light on the priestly role of the king in Crete. Religion and Power in the Ancient Greek World. *Proceedings of the Uppsala Symposium 1993*, 27–39.

Blomberg, M., & Henriksson, G. (1999). Evidence for the Minoan origins of stellar navigation in the Aegean. Actes de La Vème Conférence de La SEAC, Gdansk, 5-8 Septembre 1997, 69–81.

The Minoan peak sanctuary on Pyrgos and its context. (2003). In M. Blomberg Blomberg, P.and Henriksson, G. (Ed.), Calendars, Symbols, and Orientations: Legacies of Astronomy in Culture. Retrieved from https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-48394

Blomberg, M., & Henriksson, G. (2003). The Minoan peak sanctuary on Pyrgos and its context. In Calendars, symbols, and orientations: legacies of astronomy in

culture. *Proceedings of the 9th meeting of the European Society for Astronomy in Culture SEAC*, 127-134.

Blomberg, M., & Henriksson, G. (2005). Orientations of the late Bronze Age villa complex at Vathypetro in Crete. *Mediterranean Archaeology & Archaeometry 5*(1), 51-61.

Blomberg, M., & Henriksson, G. (2006). Orientations of the Minoan palace at Phaistos in Crete. Mediterranean Archaeology and Archaeometry. 185-192.

Blomberg, M., & Henriksson, G. (2011). The evidence from Knossos on the Minoan Calendar. *Mediterranean Archaeology and Archaeometry*, *11*(1), 59-68.

Kourtis, Spyridon & Xenakis, Apostolis & Kalovrektis, Konstantinos & Antonios, Plageras & Chalvantzi, Ioanna. (2021). An Exploratory Teaching Proposal of Greek History Independence Events based on STEAM Epistemology, Educational Robotics and Smart Learning Technologies. 120-128.

Kalovrektis, K., Dimos, I., Xenakis, A., & Kakarountas, A. (2023). "Nanotechnology as a Tool for Computational Thinking Skills using Open Hardware, Embedded Systems and Repository Platform, in Industry 4.0 Era", presented in SEEDA – CESCNM

Psycharis, S., Iatrou, P., Kalovrektis, K., Xenakis, A. (2023). The Impact of the Computational Pedagogy STEAM Model on Prospective Teachers' Computational Thinking Practices and Computational Experiment Capacities. A Case Study in a Training Program. In: Auer, M.E., Pachatz, W., Rüütmann, T. (eds.), *Learning in the Age of Digital and Green Transition*. Lecture Notes in Networks and Systems, vol 634. Springer. https://doi.org/10.1007/978-3-031-26190-9_41

Psycharis, S., Kalovrektis, K., & Xenakis, A. (2020). A Conceptual Framework for Computational Pedagogy in STEAM education: Determinants and perspectives. *Hellenic Journal of STEM Education*, 1(1), 17–32. https://doi.org/10.51724/hjstemed.v1i1.4

A. Xenakis, K. Kalovrektis, K. Theodoropoulou, A. Karampelas, G.Giannakas, D. Sotiropoulos, D. Vavougios, (2023). Using Sensors and Digital Data Collection/Analysis Technologies in K–12 Physics Education Under the STEM Perspective. In Mehmet F. T & Paula, R. L. H (Eds.), *The International Handbook of Physics Education Research: Teaching Physics* (pp. 1-42).

Linh, N.Q., & Hương, L.T. (2021). Engineering design process in STEM education: an illustration with the topic "wind energy engineers". *Journal of Physics: Conference Series*, 1835.

Tsikritsis, M., Moussas, X., & Tsikritsis, D. (2015). Astronomical and mathematical knowledge and calendars during the early Helladic Era in aegean "frying pan" vessels. *Mediterranean Archaeology and Archaeometry*. *15*, 135-149.

Tsikritsis, M., Theodossiou, E., Manimanis, V. N., Mantarakis, P., & Tsikritsis, D. (2013). A Minoan eclipse calculator, *Mediterranean Archaeology and Archaeometry*, 13, 265–276.

EUROPEAN SCIENCE EDUCATION RESEARCH ASSOCIATION

ESERA

The 15th Conference of the European Science Education Research Association (ESERA)

ESERA 2023 Cappadocia

ESERA 2023 Conference

Cappadocia, Türkiye

August 28- September 1, 2023

The 15th ESERA

EUROPEAN SCIENCE EDUCATION RESEARCH ASSOCIATION

Proceedings Book Series-I

Editors
Metin Sardag- Gokhan Kaya - Mehmet Sogut

The ESERA 2023 Conference, held in the enchanting region of Cappadocia, Turkey, from August 28 to September 1, 2023, brought together a diverse community of researchers, educators, and practitioners in science education. This proceedings series, published in five volumes, encapsulates the breadth and depth of scholarly work presented at the conference. Each volume represents a collection of select strands, showcasing cutting-edge research and innovative practices in science education. Through these proceedings, readers will gain insights into the latest trends, challenges, and advancements in the field, reflecting the dynamic nature of science education across various contexts. This series serves not only as a record of the conference but also as a valuable resource for the global science education community, fostering ongoing dialogue and collaboration.

nobel bilimsel NOBEL BİLİMSEL ESERLER
www.nobelbilimseleserler.com
nobel e-kitap nobelyayin.com
nobelkitap.com
E-KİTAP