IntechOpen

Advances in Research in STEM Education

Edited by Michail Kalogiannakis
and Maria Ampartzaki

Advances in Research in STEM Education

Edited by Michail Kalogiannakis
and Maria Ampartzaki

Published in London, United Kingdom

Advances in Research in STEM Education
http://dx.doi.org/10.5772/intechopen.98076
Edited by Michail Kalogiannakis and Maria Ampartzaki

Contributors
Caroline D. Hardin, Parama Chaudhuri, Ishin Iwasaki, Corben Roszak, Katherine LaRue, Irene Govender, Getachew Tarekegn, Jonathan Osborne, Mesfin Tadesse, Israel Kibirige, Shapule Edith Modjadji, William Deng Tap, Helicopter Mark Bulbul, Biar Simon Ajang, Zenaida Aguirre-Munoz, Brian Dang, Elias S. Loria Garro, Ann M. Gansemer-Topf, Ulricke Genschel, Xuan Hien Nguyen, Jasmine Sourwine, Yuchen Wang

© The Editor(s) and the Author(s) 2022
The rights of the editor(s) and the author(s) have been asserted in accordance with the Copyright, Designs and Patents Act 1988. All rights to the book as a whole are reserved by INTECHOPEN LIMITED. The book as a whole (compilation) cannot be reproduced, distributed or used for commercial or non-commercial purposes without INTECHOPEN LIMITED's written permission. Enquiries concerning the use of the book should be directed to INTECHOPEN LIMITED rights and permissions department (permissions@intechopen.com).
Violations are liable to prosecution under the governing Copyright Law.

Individual chapters of this publication are distributed under the terms of the Creative Commons Attribution 3.0 Unported License which permits commercial use, distribution and reproduction of the individual chapters, provided the original author(s) and source publication are appropriately acknowledged. If so indicated, certain images may not be included under the Creative Commons license. In such cases users will need to obtain permission from the license holder to reproduce the material. More details and guidelines concerning content reuse and adaptation can be found at http://www.intechopen.com/copyright-policy.html.

Notice
Statements and opinions expressed in the chapters are these of the individual contributors and not necessarily those of the editors or publisher. No responsibility is accepted for the accuracy of information contained in the published chapters. The publisher assumes no responsibility for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained in the book.

First published in London, United Kingdom, 2022 by IntechOpen
IntechOpen is the global imprint of INTECHOPEN LIMITED, registered in England and Wales,
registration number: 11086078, 5 Princes Gate Court, London, SW7 2QJ, United Kingdom

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Additional hard and PDF copies can be obtained from orders@intechopen.com

Advances in Research in STEM Education
Edited by Michail Kalogiannakis and Maria Ampartzaki
p. cm.
Print ISBN 978-1-80355-729-8
Online ISBN 978-1-80355-730-4
eBook (PDF) ISBN 978-1-80355-731-1

Meet the editors

Dr. Michail Kalogiannakisis is an assistant professor in the Department of Preschool Education, University of Crete. He graduated from the Physics Department, University of Crete, and continued his post-graduate studies at the University Paris 7-Denis Diderot where he received a DEA in Didactic Physics. He obtained a DEA and Ph.D. in Science Education from the University of Paris 5-René Descartes-Sorbonne. His research interests include science education in early childhood, science teaching and learning, e-learning, and the use of ICT in science education.

Dr. Maria Ampartzaki is an Assistant Professor of Early Childhood Education in the Department of Preschool Education, University of Crete. Her research interests include science education in the early years, ICT in education, inquiry-based and art-based learning, teachers' professional development, action research, and the pedagogy of multiliteracies, among others. Dr. Ampartzaki has a bachelor's degree in Early Years Education, an MA in Music Education, and a doctorate in Education. She has published her research in numerous academic publications and is widely referenced by educators worldwide. She has run and participated in several funded and non-funded projects regarding values education, science, and ICT in education.

Contents

Preface **XI**

Chapter 1 **1**
Research Status in Computational Thinking in STEM Education
by Irene Govender

Chapter 2 **25**
Impact of Dialogic Argumentation Pedagogy on Grade 8 Students' Epistemic Knowledge of Science
by Getachew Tarekegn, Jonathan Osborne and Mesfin Tadesse

Chapter 3 **55**
Views of South Sudanese Secondary School Teachers about the Use of Humour in the Mathematics Classroom
by William Deng Tap, Helicopter Mark Bulbul and Biar Simon Ajang

Chapter 4 **73**
Impact of Integrated Science and Mathematics Instruction on Middle School Science and Mathematics Achievement
by Zenaida Aguirre-Muñoz, Brian Dang and Elias S. Loria Garro

Chapter 5 **93**
Grade 10 Girls' Experiences in Choosing STEM Subjects in Rakwadu Circuit, South Africa
by Israel Kibirige and Shapule Edith Modjadji

Chapter 6 **113**
The Power in Groups: Using Cluster Analysis to Critically Quantify Women's STEM Enrollment
by Ann M. Gansemer-Topf, Ulrike Genschel, Xuan Hien Nguyen, Jasmine Sourwine and Yuchen Wang

Chapter 7 **133**
Robots, Everlasting? A Framework for Classifying CS Educational Robots
by Ishin Iwasaki, Corben Roszak, Parama Chaudhuri, Katherine LaRue and Caroline D. Hardin

Preface

Science, Technology, Engineering, and Mathematics (STEM) education is considered an educational and political priority almost all over the world. It is an interdisciplinary methodology that teaches STEM subjects through a combination of theoretical academic concepts and real-world applications. STEM education is inspired by the theoretical framework of Social Constructivism where learning is not approached as an individualised process but is instead considered to take place in a social context.

The global education status quo acknowledges STEM education as being a high priority. Many countries have revised their education systems and focused on new educational methodologies to equip students with the necessary skills and match them with the demands of the competitive job market [1]. Future employees need STEM skills to be able to cope with work demands and thus STEM education is essential. Many international initiatives consider STEM education as one of the best academic investments. It fosters broad-based scientific literacy to improve undergraduate education in elementary, middle, and secondary schools. In addition, science and technology have a critical role in achieving the United Nations' Sustainable Development Goals (SDGs) [2] and ensuring the prosperity of all people [3].

Over the years, STEM evolved into a pedagogical approach that includes the four subjects/disciplines, with the fundamental goal of presenting academic concepts in realistic and meaningful situations. Students are called to solve real-world problems using the research process, as professional scientists do [4]. In various countries, the development of STEM skills is a high priority and an essential prerequisite to innovations or the competence that stands out and prevails in the international job market. STEM-focused schools have been created, emphasizing technology in long-term learning experiences in specific projects while promoting critical thinking, communication, and collaboration. To address the ongoing need for more STEM-literate workers, classrooms at all educational levels are integrating STEM curricula [5]. However, successful integration and implementation of STEM curricula is challenging, mainly because educators must develop a comprehensive understanding of the integration concepts, strategies, and tools [6–8]. In addition, technology brings rapid advances in education with new educational kits and methods that enable teachers to design learning content adapted to the demanding needs of the students. This can present a challenge, as it requires a great effort for preschool and elementary teachers who often have limited knowledge of STEM content [9].

In the last decade, there has been a wave of enthusiasm for STEM education to improve early childhood education programs. Research has indicated that it would be helpful for children's development if children could be exposed to STEM education at a young age [10]. Is it essential to start STEM education at early school age? De Vries et al. [11] indicate that science and technology are important in young students' education, and studies suggest that science and technology can help young students to:

- keep up with the rapid and continuous changes that technology and science bring,
- understand and recognize science as an essential human achievement, and
- know how to approach problems by looking for relevant information and making evidence-based decisions.

However, even in countries where STEM education has been generally recognized and valued for several years now, an integrated approach is mainly implemented in secondary education. This creates many questions about whether early exposure to STEM would positively affect young students. Research has shown that this would be possible by providing a developmentally appropriate STEM education through authentic classroom experiences, which will familiarize students with the nature of professional work in STEM and what STEM professionals really do [12, 13]. Moreover, research reveals gender gaps in STEM education, as fewer girls and women choose to study and pursue STEM careers [14, 15].

Advances in Research in STEM Education edited by Michail Kalogiannakis and Maria Ampartzaki showcases some recent developments in STEM education.

Chapter 1, “Research Status in Computational Thinking in STEM Education”, gives a global overview of the status of computational thinking in STEM and related fields. Because of its importance, particularly in this era of the Fourth Industrial Revolution (4IR), this chapter focuses on an awareness of how the teaching thereof and its implementation in different areas are reviewed. The inherent value of computational thinking in the STEM disciplines cannot be underemphasized. A systematic literature review was the basis for this chapter. While it has been shown in other chapters that there is a need to attract and retain students in STEM fields, it is noted that computational thinking must consciously be enacted within the STEM and related fields to meet the demands of the 4IR. This chapter targets all stakeholders (educators, researchers, and academics) in STEM education.

Chapter 2, “Impact of Dialogic Argumentation Pedagogy on Grade 8 Students’ Epistemic Knowledge of Science”, explores the effect of dialogic argumentation on Grade 8 students’ epistemic knowledge of physics. The authors employ a mixed-methods experimental design research approach, using a quasi-experimental design to compare the epistemic knowledge of science in argumentation lessons between an experimental and a control group. The study found that argumentation lessons significantly increased students’ level of epistemic knowledge compared to students who did not practice argumentation. In addition, the lessons were found to significantly improve the quality of students’ argumentation as well.

Chapter 3, “Views of South Sudanese Secondary School Teachers about the Use of Humour in the Mathematics Classroom”, explores the views of secondary school teachers about the strategic use of humour in the mathematics classroom as a teaching and learning technique or strategy. This is to inspire, motivate, generate, and maintain students’ interest in mathematics, one of the key STEM subjects. The overall aim of the chapter is less about the introduction and promotion of the specific use of humour in the math classroom and more about presenting, advocating, and arguing for a diverse and holistic approach to the teaching and learning of mathematics. Mathematics and its other STEM-related disciplines are generally challenging to many students and therefore it pays to enlighten the subject in ways that are appreciated by the learners. The use of humour by a teacher in classroom settings is one technique that points in this direction. Hence, although the use of humour in the class (often referred to in the literature as classroom humour or instructional humour) is just one technique out of many, it may be the key to the elusive concept of "best teaching practices" or "instructional best practices" in classroom settings.

Chapter 4, "Impact of Integrated Science and Mathematics Instruction on Middle School Science and Mathematics Achievement", documents the impact of the Middle School Math and Science (MS)2 Integration project, employing the results of the internal evaluation of this intensive teacher training model for integrated science and mathematics in middle school. The authors mention that although students in (MS)2 classrooms are more likely to have higher achievement scores, the frequency of integrated instruction opportunities also significantly predicts student achievement, particularly in mathematics classrooms.

Chapter 5, "Grade 10 Girls' Experiences in Choosing STEM Subjects in Rakwadu Circuit, South Africa", acknowledges that few women pursue STEM careers, although all learners in South Africa from Grades 5 to 9 study mathematics and general science, which combines physics, chemistry, biology, and earth science. However, fewer girls choose to study science subjects in Grade 10. Schools have a good balance of boys and girls studying general science subjects because they are compulsory until Grade 9. In Grade 10, learners choose to study arts, business, or STEM subjects. It is at this level that the number of girls in STEM subjects drops dramatically, leaving the boys to continue unabated, sustaining the hegemony of the masculine gender in STEM subjects. Hence, the study seeks to understand girls' experiences in choosing STEM, to assist education stakeholders in motivating and advocating for more girls to choose science subjects and STEM careers. The case study in this chapter focuses on ten girls who decided to study science out of 145 girls who attended general science in Grade 9. The study reveals that personal factors, anticipated value, the class environment, home influence, and social influence contributed significantly to the girls' choices. If such a drop from Grades 9 to 10 continues, the hope of decreasing the gender gap in STEM will be lost. The increase of girls in STEM subjects at the secondary school level can, in the long run, narrow the gender gap in STEM careers.

Chapter 6, "The Power in Groups: Using Cluster Analysis to Critically Quantify Women's STEM Enrollment", argues that despite efforts to close the gender gap in STEM, disparities still exist, especially in math-intensive STEM (MISTEM) majors. Females and males receive similar academic support and, overall, perform similarly, yet females continue to enroll in STEM majors less frequently than men. In examining academic preparation, most research considers performance measures individually, ignoring the possible existence of interrelationships between these measures. The authors address this problem by using hierarchical agglomerative clustering, a statistical technique that allows for identifying groups (i.e., clusters) of students who are similar in multiple factors. They apply this technique to readily available institutional data to determine if distinct groups could be identified. Results detected nine unique groups. Following this, the authors explored differences in STEM enrollment by group and by gender. The outcomes showed that the proportion of females differed by group, and the gap between males and females also varied. Overall, males enrolled in STEM at a higher proportion than females, regardless of the strength of their academic preparation. This chapter presents a novel yet practicable and realistic approach to examining gender differences in STEM enrollment in postsecondary education.

Chapter 7, “Robots, Everlasting? A Framework for Classifying CS Educational Robots”, explores the history of issues in device purchases, documents several examples of equipment breakdown, and details the unique and specific needs of school customers. The chapter introduces the Computer Science Risk Analysis Framework for Toys to help teachers and schools evaluate a device purchase based on a holistic understanding of device longevity. This study also provides recommendations for computer science (CS) and STEM educational robot designers.

Apart from being a practical and helpful resource for curriculum innovation projects *Advances in Research in STEM Education* provides thought-provoking information for educators, education leaders, education researchers, and stakeholders in early childhood, primary, and higher education institutions. It contains theoretical and pedagogical frameworks, trends, samples of good practices, and a discussion of challenges. Directions and implications for future research identified by the studies in this book will help the audience to gain a comprehensive and deeper understanding of STEM education.

Michail Kalogiannakis and Maria Ampartzaki
Department of Preschool Education,
University of Crete,
Rethymno, Greece

References

[1] Barak, M., & Assal, M. (2016). Robotics and STEM learning: Students' achievements in assignments according to the P3 Task Taxonomy - practice, problem solving, and projects. *International Journal of Technology and Design Education*, *28*(1), 121–144, https://doi.org/10.1007/s10798-016-9385-9

[2] United Nations. (2015). *Sustainable Development Goals | United Nations Development Programme*. UNDP. Retrieved 5 September 2022, from https://www.undp.org/sustainable-development-goals

[3] Vuong, Q. H. (2020). *From Children's Literature to Sustainability Science, and Young Scientists for a More Sustainable Earth* (SSRN Scholarly Paper No. 3754465).

[4] Williams, J. (2011). STEM education: Proceed with caution. *Journal of Design and Technology Education*, *16*(1), 26–35.

[5] Margot, K. C., & Kettler, T. (2019). Teachers' perception of STEM integration and education: A systematic literature review. *International Journal of STEM Education*, *6*(2), https://doi.org/10.1186/s40594-018-0151-2

[6] Psycharis, S. (2018). STEAM in education: A literature review on the role of computational thinking, engineering epistemology, and computational science. Computational STEAM Pedagogy (CSP). *Scientific Culture*, *4*(2), 51–72.

[7] Ampartzaki, M., Kalogiannakis, M., Papadakis, St., & Giannakou, V. (2022). Perceptions About STEM and the Arts: Teachers', Parents' Professionals' and Artists' Understandings About the Role of Arts in STEM Education. In St. Papadakis & M. Kalogiannakis (Eds), *STEM, Robotics, Mobile Apps in Early Childhood and Primary Education - Technology to promote teaching and learning.* Lecture Notes in Educational Technology, (pp. 601-624). Switzerland, Cham: Springer, https://doi.org/10.1007/978-981-19-0568-1_25

[8] Kalogiannakis, M., & Papadakis, S. (2022). Preparing Greek Pre-service Kindergarten Teachers to Promote Creativity: Opportunities Using Scratch and Makey Makey. In K.-J, Murcia, C., Campbell, M.-M. Joubert & S. Wilson (Eds.), *Children's Creative Inquiry in STEM. Sociocultural Explorations of Science Education*, vol 25. (pp. 347-354), Switzerland, Cham: Springer, https://doi.org/10.1007/978-3-030-94724-8_20

[9] Estapa, A. T., & Tank, K. M. (2017). Supporting integrated STEM in the elementary classroom: a professional development approach centered on an engineering design challenge. *International Journal of STEM education*, *4*(1), 6, https://doi.org/10.1186/s40594-017-0058-3

[10] Kazakoff, E. R., Sullivan, A., & Bers, M. U. (2013). The effect of a classroom-based intensive robotics and programming workshop on sequencing ability in early childhood. *Early Childhood Education Journal*, *41*(4), 245-255, https://doi.org/10.1007/s10643-012-0554-5

[11] De Vries, M. J., van Keulen, H., Peters, S., & van der Molen, J. W. (Eds.). (2012). *Professional development for primary teachers in science and technology* (Vol. 9). Springer Science & Business Media.

[12] DeJarnette, N. (2012). America's children: Providing early exposure to STEM (science, technology, engineering and math) initiatives. *Education*, *133*(1), 77-84.

[13] Murphy, T. P., & Mancini-Samuelson, G. J. (2012). Graduating STEM competent and confident teachers: The creation of a STEM certificate for elementary education majors. *Journal of College Science Teaching*, *42*(2), 18–23.

[14] Bowman, N. A., Logel, C., LaCosse, J., Jarratt, L., Canning, E. A., Emerson, K. T. U., & Murphy, M. C. (2022). Gender representation and academic achievement among STEM-interested students in college STEM courses. *Journal of Research in Science Teaching*, *n/a*(n/a), 1–25. https://doi.org/10.1002/tea.21778

[15] Stoet, G., & Geary, D. C. (2018). The Gender-Equality Paradox in Science, Technology, Engineering, and Mathematics Education. *Psychological Science*, *29*(4), 581–593, https://doi.org/10.1177/0956797617741719

Chapter 1

Research Status in Computational Thinking in STEM Education

Irene Govender

Abstract

Computational thinking (CT) is an approach to problem-solving that has its roots in computer science. However, its inherent value in the science, technology, engineering, and mathematics (STEM) disciplines cannot be over-emphasized, considering that we are in the fourth industrial revolution. The chapter draws attention to its close affinity to problem-solving and programming, and the impact of computational thinking on the labour market, and in turn the digital economy is highlighted. A global overview of recent research findings and initiatives to implement CT education in school curricula are discussed. Because of the importance of STEM education, and the inherent value of CT, it is necessary to explore the status and inclinations of CT in STEM disciplines. Hence, a snapshot of research over the last two years was used in a systematic review to determine the trends and challenges for integrating CT in the curriculum of STEM related fields. Using the ERIC database of journals, and specific criteria for selection of publications, 31 articles were examined in this study. Overall, it was found several tools and instructional strategies are used to develop CT, but more needs to be done to increase teachers' knowledge and enactment for CT in the STEM fields.

Keywords: STEM, computational thinking, problem-solving, artificial intelligence, teachers, programming, robotics

1. Introduction

Computational Thinking is fundamental for many, if not all occupations, particularly science, technology, engineering, and mathematics (STEM). STEM related fields play a significant role in economies by driving innovation to meet the demands of the fourth industrial revolution era. However, STEM fields of study are often perceived as difficult and many students drop out of these subject areas as a result, impeding career opportunities in the related fields [1]. Accordingly, institutions of higher learning play a crucial role in preparing the people for STEM employment to meet the exigences of the 4th industrial revolution. While the acronym, STEM, was coined as a general and appropriate word for Science, Technology, Engineering and Mathematics fields of study, STEM often relate to all sciences (astronomy, physics, computing fields and the like). These fields often depend on computational tools for modeling and simulation, data analysis and visualization, creating computer programs to solve problems, and understanding a system as an aggregate of parts;

these are characteristics of computational thinking. Hence, it can be inferred that Computational thinking (CT) is inherent to STEM practices [2]. Moreover, computational thinking is widely considered to be an important and necessary twenty-first century skill that contributes to the development of solving complex problems. As a result, a growing body of literature has investigated the tools and interventions to foster and develop CT in education [3–5]. Overall, such studies highlight the need for a consistent outcome of the interventions implemented.

Moreover, with the growing importance of STEM education world-wide [6], it is not surprising for the need to foreground research, not just in general STEM education, but in the integration of computational thinking in the STEM fields. The need for CT to be enhanced and fostered in STEM fields is imperative. Therefore, this chapter seeks to assess the research status and inclinations of computational thinking in STEM fields and determines the key findings for future implications.

The rest of the chapter proceeds as follows: a) Literature review highlighting the key concepts of STEM, CT and its close affinity to programming and AI, b) Methodology), c) Results, d) Discussion of findings, and d) Conclusion. This chapter contextualises the research by providing background literature on CT, programming and AI, and the global status of CT education. The chapter then discusses the specific methods by which the research and analyses were conducted, followed by a discussion and conclusion section.

2. Literature review

2.1 What is computational thinking

It is important to re-iterate that while computational thinking has its roots in computer science (CS), it is not computer thinking and reasoning, or programming either. Though there is minimal consensus on the definition of computational thinking, there is agreement in the literature that CT involves decomposition, abstraction, pattern recognition, and algorithmic thinking, which when expanded entails the following:

Decomposition: Splitting a composite problem or system into smaller components and solving each component and then logically organizing and analyzing data and making deductions.

Abstraction: Managing complexity so that the complicated and difficult aspects can be put aside into a black box so others can work on the details of it, focusing on the relevant information only, while temporarily ignoring the detail in the black box. Abstraction is at the heart of CT.

Pattern recognition: Looking at how people have solved similar problems drawing on that experience and identify similarities among and within problems.

Algorithmic thinking: Formulating a set of steps to achieve the objective, i.e., a set of instructions.

Generalizing this problem-solving process: Translating trends and patterns into rules, principles, or insights to apply to a wide variety of problems.

After igniting the importance of computational thinking advocacy in 2006, ten years later Jeanette Wing [7] defined CT as

> *"... the thought processes involved in formulating problems and their solutions so that the solutions are represented in a form that can be effectively carried out by an information-processing agent" (p. 8).*

On examining this definition, two aspects emerge that are of importance for STEM education.

Computational Thinking (CT) is a thinking procedure; therefore, CT can be independent from the use of technology. Supporting this view, Sysło [8] writes that CT "is a collection of key mental tools and practices originated in computing but addressed to all areas far beyond computer science" (p. 1). Furthermore, there is strong evidence that unplugged approaches can be just as effective, if not better in advancing computational thinking skills and thereby facilitating students' ability to program [3]. For example, in their study [9] found that when comparing plugged to unplugged (non-computer use) approaches as students learned programming, those who learned using unplugged approaches were more confident in understanding the concepts than those who used the plugged approach. Hence, it can be inferred that CT can be used in other STEM fields, where the use of computers is not required. It is, therefore, not surprising that [4] in an earlier study described computational thinking as, "… a specific type of problem-solving that entails distinct abilities, e.g., being able to design solutions that can be executed by a computer, a human, or a combination of both". Thus, to the non-specialist in computing, good sense suggests that computational thinking may be explained as a way of reasoning and of solving problems in a modern-day world characterized by up-to-date technology.

2.2 Importance of computational thinking (CT)

While the genesis of computational thinking can be traced back to Papert [5] for his work in Logo programming, it was only when Wing [10], the former Vice President of Microsoft Research, published her seminal paper on CT, did research in CT begin to gain momentum. As we have been ushered into the 4th Industrial revolution, this increased attention to CT could not have been timelier. In the last decade, there has been a surge of interest in the effects of computational thinking. While teaching children thinking is a key competence that education should inculcate, developing CT has become even more crucial in this era of digital economy.

The advancement of digital technology has heightened the use of computational thinking and this trend is growing across all industries, which certainly has implications for our students and the labour market. This situation in turn should impact our education and STEM curriculum.

Hence, solving unusual problems in the current era of digital technology is an important competence. Our current students live a life greatly predisposed to information technology (IT), and many will work in areas that involve or are impacted by IT. The unprecedented advancement of technology, and its various forms of communication via the Internet, have not only permeated our lives in many respects, but is hugely impacting the digital economy. To name a few examples; in healthcare – operating rooms' efficiency depends on computing, and it has enabled advances and inventions such as contact lenses that detect levels of insulin in people with diabetes; in space – there is a move to develop and use a generation of robots to explore where humans cannot now; in households – people have begun to automate every-day phenomena like the heating and lighting systems, and the use of robots to clean floors and carpets; on our road travels – we depend on navigation systems to get us to our various destinations, and now manufacturers are experimenting with self-driving cars. Hence, computational thinking has been recognised as a twenty-first century skill [6, 11].

Wing [7] in a later study asserts that:

> *"Everyone should be given the opportunity to gain competences in the field of computational thinking to allow them to successfully participate in a digitalized world".*

This excerpt gives importance to the claim that computational thinking is fundamental, not only for computer science but for all sciences in agreement with [12]. This claim raises new charges and challenges for schools. Furthermore, the increasing rate of technology users or consumers far exceeds that of creators or digital innovators disproportionately. This imbalance can have some adverse effects on the labour market and consequently the digital economy. How do we close this disproportionate gap? This situation, therefore, calls urgent attention to the development of computational thinking generally, and more specifically in STEM fields. While much research has been generated in computational thinking and in STEM education [1], comparatively there are limited studies on the integration of computational thinking and its use in STEM fields. The question therefore arises: how do we know whether our students really possess this skill to meet the twenty-first century skills set? Moreover, with the rise of artificial intelligence awareness, tools, and applications, computational thinking is even more crucial in this era. Several studies advocate programming and AI for all, inferring that coding, or programming is needed in most STEM related fields. What follows, is a discussion of the link between CT, programming, and AI.

2.3 Computational thinking, programming and artificial intelligence

Computing is at its heart a creative subject. The best way to learn is to make something by getting involved. Children learn by playing and experimenting with technology, in keeping with Seymour Papert's theory of learning—Constructionism. Several efforts to develop students' CT has tended to use activities or tools that are directly linked to programming skills in educational settings (e.g. [13–15]). However, this close kinship to computer science or IT, does not automatically imply that computational thinking is exclusively the domain of computer science or programming. Still, it is not injudicious to consider the relationship between CT and programming.

As mentioned earlier, one of the aspects of CT is algorithmic thinking. An algorithm is an unambiguous defined step by step guide to solve a problem or achieve a particular objective. Programming, however, may be broadly regarded as a two-step process – first, a set of steps to solve a problem, its algorithm and second, the coding of the algorithm into a specific system. How one solves the problem to achieve the solution involves the creative thinking. One then has the task of making those thoughts come into action by translating that algorithm into a set of symbols according to the computation system one is using, i.e., the language. Both are hard and both require creativity.

So, the best way to learn is by doing and the practical experience of programming [is] almost certainly the best way for primary school pupils to learn about computing and Computational thinking. This has been observed anecdotally and from the literature that is beginning to show evidence of this trend.

More recently, Computational Thinking (CT) and programming skills are now deemed as being as essential as numeracy and literacy by many scholars [16]. In short, getting computers to help us to solve problems is a two-step process. First, we think about the steps to solve a problem or the rules that govern a system. Second, we use our technical skills to get the computer working on the problem. Computational Thinking is the first of these. It describes the concepts, processes, and approaches we draw on, when thinking about problems or systems in a way that a computer can help us with these. It is really that first stage that we ought to be concentrating on and it is very much of what we do in computing. For example, if one is going to write up something – one thinks of the idea, structure, the content, etc. before one starts to type it up.

Having discussed the association between CT and programming, I will now move on to discuss their association to AI, a technological revolution in terms of innovation.

Heintz [17], a specialist in AI, indicated in his talk that "Computational thinking develops techniques for people to solve problems in a way that allows computers to help. Artificial intelligence develops techniques for computers to solve problems." CT captures what we need to be good at to leverage all the artificial intelligence (AI) and other computational tools that are available.

What is interesting here is that as we learn more about AI and as these tools are developed based on for example machine learning we need to be able to leverage them through our CT skills – by being able to understand how the computational processes work and how we can benefit from those when we solve problems. There are several interesting cases where people use machine learning or other tools as part of their own problem-solving process.

Artificial intelligence platforms involve the use of machines to mimic human reasoning. In an attempt to mimic human cognition, these platforms model human reasoning, problem-solving and intelligence, both social and general of which computational thinking is part.

Referring to Papert's [5] paper, he said "...technology is something children themselves will learn to manipulate, extend, to apply to projects, thereby gaining a greater and more articulate mastery of the world, a sense of the power of applied knowledge" (p 353). Using the power of computing to make a meaningful social impact, children are empowered to make an impact by doing activities situated in context that are personally relevant. Hence, in our world of AI, computational thinking skills should be a core competence for all students.

Returning to the subject of computational thinking, in the section that follows, I review and summarize the global status of CT education.

2.4 Global status of computational thinking (CT) education

There has been much research on interventions to include CT in the curriculum – these interventions invariably involve some aspects of programming. However, in his paper, Yadav [18] foregrounds how CT nudges students past operational and technical skills, creating problem solvers instead of consumers of software and technology. In a later study, [19] assert that because of its capacity for automation and enactment, programming appears to be a natural vehicle for learning computational thinking, but with some amendments in the approach to focus on CT.

A noteworthy finding in Taslibeyaz, Kursun, and Karaman's [20] study is that the development of CT skills is predominantly examined with programming content tools, such as Scratch and robotics for school level students. Additionally, it was found that studies on the development of CT for university students were carried out in other content areas besides programming education. While some studies indicated that teaching CT does improve programming education, others have found that programming courses develop CT.

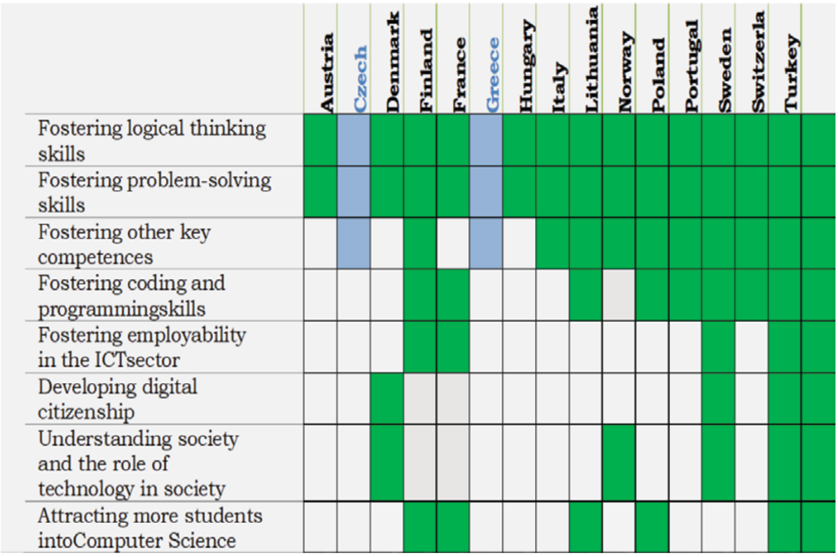
Ministries of education in many countries have recognized the importance of computational thinking for their economy [16]. In a joint report compiled by JCR, it was envisaged that both CT and programming are key competencies for compulsory education. This report indicates the countries that have recognised the importance of CT and its implementation in the school curricular. While **Table 1** is not meant to be exhaustive or comprehensive, it provides an overview of the countries that have implemented CT as compulsory education in the school curriculum at the time of writing this chapter. The piloting and revision of the implementation of CT has been done in the years prior to what is indicated.

England (part of UK) was one of the first countries to include CT as part of the mandatory course in the school curriculum as early as 2014. Interestingly, by 2018 all 50 states in the US had implemented CT education in the school curriculum as policy. In countries like Ghana and Nigeria – there is commitment to include CT as a mandatory aspect –but it has not yet become policy. In India – in the rural areas Computer science (CS) has become compulsory – they have worked with CSpathshala, an ACM India initiative to bring a computing curriculum to Indian schools in 11 different states. At the time of writing this chapter, it was determined that they are teaching it in the mathematic curriculum, while there is a huge drive to include all schools.

While South Africa has the commitment to implement CT in the school curricula – they have not yet successfully piloted the implementation, which was due to have started in 2021 in 1000 schools. It is necessary to note the huge challenges of

Country (not meant to be exhaustive)	Implement Year	Current state of new policy initiatives
Asian countries (Taiwan, Japan, and Korea)	2020	For primary in 2020 and -2022 for secondary
Australia	2018	Digital technologies compulsory with CT education
European countries (17 countries—Austria, Czech Rep Finland, Denmark, France, Greece, Hungary, Ireland, Italy Lithuania, Norway, Poland, Portugal, Russia, Sweden, Switzerland, Turkey)	2018	Denmark and Norway introducing CT education as a permanent elective. Piloting began in 2016
UK (England, Scotland, Netherlands, Wales)	2018	England is the first country to make CT education compulsory in 2014
USA	2018	By 2018 all 50 states have policy for CT in schools
Hong Kong	2017	CT supplement introduced to Primary from P4 to P6
New Zealand	2017	The digital curriculum was reformed to include CT
Singapore	2014	CT implemented as optional in different education sectors—called smart Nation initiative
Ghana		Part of a code club – commitment to implement CT- no formal policy
Nigeria		no formal policy- focus on enrichment
India		Advocacy phase- CSpathshala initiative
South Africa		A Pilot was planned for 2021.

Table 1.
Global status of CT education.



	Austria	Czech	Denmark	Finland	France	Greece	Hungary	Italy	Lithuania	Norway	Poland	Portugal	Sweden	Switzerla	Turkey	
Fostering logical thinking skills	🟩	🟦	🟩	🟩	🟩	🟦	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
Fostering problem-solving skills	🟩	🟦	🟩	🟩	🟩	🟦	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
Fostering other key competences		🟦		🟩		🟦		🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
Fostering coding and programmingskills				🟩	🟩				🟩	▒	🟩	🟩	🟩	🟩	🟩	🟩
Fostering employability in the ICTsector				🟩	🟩								🟩		🟩	🟩
Developing digital citizenship			🟩	▒	▒								🟩		🟩	🟩
Understanding society and the role of technology in society			🟩	▒	▒					🟩			🟩		🟩	🟩
Attracting more students intoComputer Science				🟩	🟩				🟩		🟩				🟩	🟩

Figure 1.
Rationale for integrating CT in the curriculum (adapted from [16]).

infrastructure and teacher development knowledge that exists. It is envisaged that universities as stakeholders for preparing students for the workplace, have a role to play in addressing these challenges.

In their report, Bocconi, et al. [4] review the ministries of education of several countries regarding the status of CT in the school curriculum. In attempting to embed CT, CS, or coding in the school curriculum, **Figure 1** summarizes the rationale for including CT in the curricula across the globe. It was found that most countries introduced CT in mainstream education at secondary level [4]. However, an emerging trend indicates integration of CT in primary school levels.

What is important to recognize here, is that fostering logical thinking and problem-solving skills are the two most common reasons for including CT in the curriculum, followed by other competences and programming. From the literature, it can be summarized that the three main reasons for including CT in compulsory education are: developing CT to increase economic growth, occupy ICT related vacancies, and groom for future work or occupations [9].

To further determine the status of Computational thinking related to studies in STEM education – the fields that drive the economy, a systematic review methodology was employed to ascertain the trends of these studies. The details of the review are presented in the Methodology section. I synthesized this review with existing ones to inform the global status of CT in STEM education and the empirical evidence of CT development.

3. Methodology

Following the PRISMA model (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [21], a systematic literature review was conducted. In the current study, the following steps were taken to achieve the relevant set of articles for analysis.

1. Specifying the research questions to guide the search.

2. Determining the database(s) for the search procedure
3. Specifying the inclusion and exclusion criteria
4. Choosing the studies
5. Conducting a pre-analysis and extraction of data

3.1 The research questions

To guide this review research the following three research questions were formulated:

- What is the status of CT in STEM education research from 2020 to the end of 2021 based on refereed journal publications?
- What research methods did authors tend to use in conducting CT in STEM education research?
- What key themes had emerged in CT in STEM education research based on the journal publications?

3.2 Determining the database(s) for the relevant studies

The review focused on publications that appeared beyond specific discipline-based journals. Due to the plethora of studies generated in the last few years on the specific aspects of CT and STEM separately, it was decided to examine the most recent research on this exciting and relevant field during the pandemic, CT in STEM fields. It was therefore prudent to consider a snapshot of studies that was extracted from January 2020 until December 2021.

It was assumed that articles on CT in STEM education have been published in journals that involve more than one conventional discipline. Since there are too many conventional discipline-based education journals, journals were not selected but emerged as part of the search results. Since this review is embedded in the education context, it was considered viable to look at the ERIC (Education Resources Information Center) database of journals. The ERIC database is a comprehensive database with information and studies in all disciplines related to education, consisting of several education journals. Using the EBSCOhost Research Databases interface, the advanced search option was selected to include the database: ERIC. The automation search strategy provided, enabled me to expand the search and to apply limitations to suit the study. The results of the search yielded source types from 76 academic journals and 76 reports.

3.3 Exclusion and inclusion criteria used

It was necessary to select studies that would be appropriate to the goal of this review, based on the review questions. The following criteria were used in selecting the studies:

3.3.1 Inclusion criteria

- Studies must be empirical (quantitative, qualitative, or mixed methods) in an educational environment.

- Any STEM related subjects studied with CT
- The search was expanded to apply related words and equivalent subjects.
- The article is a peer-reviewed study
- The momentum of studies generated regarding CT and studies regarding STEM have increased exponentially. Hence, the scoped articles were from 2020 until 2021.

3.3.2 Exclusion criteria

- Studies written in a non-English language.
- Studies only published as an abstract
- Conference papers—Since journal publications are acknowledged as one of the quality sources of research ideas and outputs (e.g., [22, 23]), articles published only in journals were considered, all other publications, including conference proceedings and grey matter were excluded.

3.4 Identifying articles

As was pointed out in the introduction to this paper, the acronym STEM relates to all sciences relating to the core subjects of science, technology, engineering, and Mathematics, which are being recognized as a global interdisciplinary field for our students to learn. Moreover, computational thinking has also been recognised as a twenty-first century skill in the current 4th industrial revolution. Using the phrase "computational thinking" together with STEM related fields as identifiers following the methodology of other researchers [24, 25], a set of relevant research articles were obtained. Specifically, the Boolean phrase used was "Computational Thinking AND (mathematics or computer science or engineering or technology or STEM)".

3.5 Pre-analysis and extraction of data

Additional criteria were imposed in the present literature research: the article abstracts were screened for empirical interventions and outcomes related to CT.

Based on relevance to the research questions, 30 articles were excluded. The remaining 46 were scoped for further information. Studies that did not include the specific interventions to develop CT in any STEM field education were excluded from the review. The remaining 31 articles were then examined in detail in relevance to the criteria and research questions. Each article was read twice to note and understand the content, procedures, and methods used, and outcomes. The PRISMA process that I followed is depicted in **Figure 2**.

The 31 articles that composed the final dataset were included in the systematic review as shown in **Table 2**. The studies were examined for contexts, content areas or interventions, variables, and their relationships with each other.

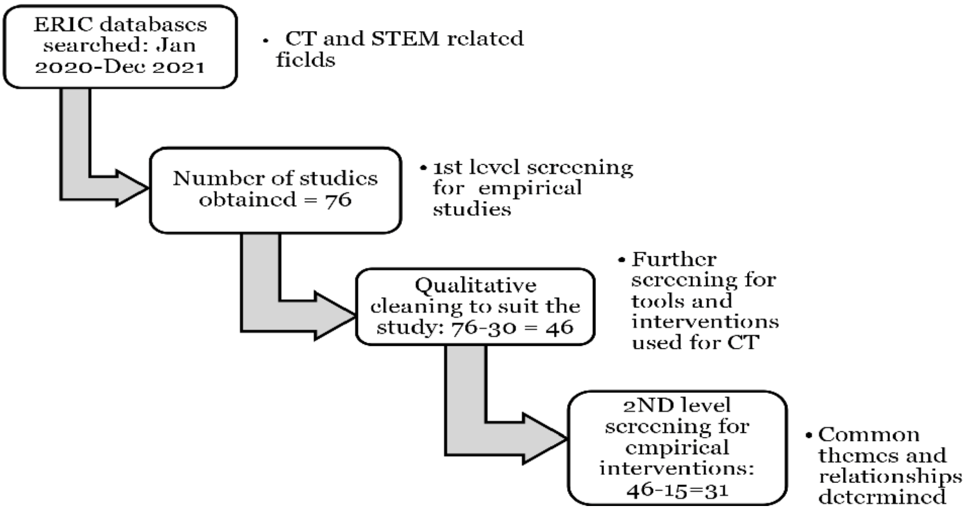


Figure 2.
Process used to obtain studies.

SN	Author(s) and year	Context/subject	level	Approach
1	Hébert and Jenson (2020) [26]	science	Secondary	qualitative
2	Lyon and Magana (2021) [27]	Engineering	Undergraduate/ tertiary	qualitative
3	Ardito, Czerkawski and Scollins (2020) [28]	Programming/ engineering	Primary	Qualitative
4	Zha, Jin and Moore (2020) [29]	Programming	Pre-service teachers/ tertiary	mixed method
5	Kynigos and Grizioti (2020) [30]	programming/ gaming	Secondary	qualitative
6	Hunsaker and West (2020) [31]	Interdisciplinary project—CT and coding	Preservice/tertiary	qualitative
7	Deniz, Kaya and Yesilyurt (2021) [32]	integrated stem	Primary/secondary	qualitative
8	Lapawi and Husnin (2020) [33]	Science module	Secondary	quantitative
9	Kukul and Çakir (2020) [34]	programming	primary, undergraduate / tertiary	qualitative
10	Çevik, Baris and Sirin (2021) [35]	Science		mixed method
11	Ilic (2021) [36]	Technologies course	pre-service teachers/ tertiary	mixed method
12	Pürbudak and Usta (2021) [37]	Information Technology	Primary	quantitative

SN	Author(s) and year	Context/subject	level	Approach
13	Ntourou, Kalogiannakis and Psycharis (2021) [38]	Science	Primary	quantitative
14	Yildiz and Seferoglu (2021) [39]	programming	Secondary	quantitative
15	Usengül and Bahçeci (2020) [40]	programming / science	Primary	quantitative
16	Avcu and Er (2020) [41]	programming	Primary	quantitative
17	Emara, Hutchins and Grover (2021) [42]	science, computing	Secondary	qualitative
18	Eryilmaz and Deniz (2021) [43]	programming	Secondary	mixed method
19	Karakasis and Xinogalos (2020) [44]	programming	Teachers	qualitative
20	Türker and Pala (2020) [45]	Programming-Computer education	pre-service/tertiary	mixed method
21	Min and Kim (2020) [46]	Computing software	Primary	qualitative
22	Threekunprapam and Yasri (2020) [47]	programming	Secondary	qualitative
23	Tsakeni (2021) [48]	science	pre-service teachers/ tertiary	qualitative
24	Delal and Oner (2020) [49]	Computing software	Primary/secondary	qualitative
25	Chongo, Osmanand Nayan (2021) [50]	Chemistry	Secondary	quantitative
26	Hijón Neira, García-Iruela and Connolly (2021) [51]	programming	Secondary	qualitative
27	Robertson, Gray, Martin and Booth (2020) [52]	programming	Primary	qualitative
28	Kopcha, Ocak, and Qian (2021) [53]		Primary	qualitative
29	Noh and Lee (2020) [54]	programming	Primary	mixed method
30	Avcu and Ayverdi (2020) [55]		Secondary	quantitative
31	Chongo, Osman and Nayan (2020) [56]	STEM	Secondary	quantitative

Table 2.
Articles included in the review.

3.6 Data analysis

To address the research questions, first the keywords were examined to determine the most common keywords and its importance to the research studies. Other descriptive statistics were determined. To address research question 2, all 31 identified publications were examined for the approaches used (1) qualitative, (2) quantitative, (3) mixed methods, and (4) non-empirical studies (including theoretical or conceptual papers, and literature reviews).

Thereafter, the themes related to the interventions or tools were identified and used in the review of publications identifying intervention tools and its relationship to the variables identified.

4. Results

4.1 Initial findings and descriptive analyses

To obtain an overview of the main themes researched, a word cloud of keywords of the 31 articles was performed. The top five keywords that emerged were computational thinking, programming, coding, learning skills, and education (**Figure 3**). These initial findings inferred the main aspects of CT in STEM are the education of CT and its relation to programming and coding.

Regarding the participants in the studies reviewed, it can be seen from **Table 2**, that there were more studies undertaken with primary (14) and secondary (14) students than tertiary (6) students or teachers in service (1). In some cases, the participants covered all three groups of participants, giving a sum of studies more than the total number of articles reviewed. This may be related to the need to develop school pupils' CT skills early before joining the job market or entering institutions of higher learning. Regarding university students (tertiary), they have been involved more as research subjects in STEM subjects and most studies involved teacher education students to determine their knowledge of CT to develop their pedagogical practices. These studies focused on educational implementation of computational thinking that was central to the study. Considering **Table 2**, it shows that the studies were carried out in the most common context, programming (F=15) and in all other STEM related fields (F= 16).

4.2 Approach used in the data Analyses

This section examines the approaches used to analyze data in the 31 studies that were included in the systematic review. Most of the studies were found to use a

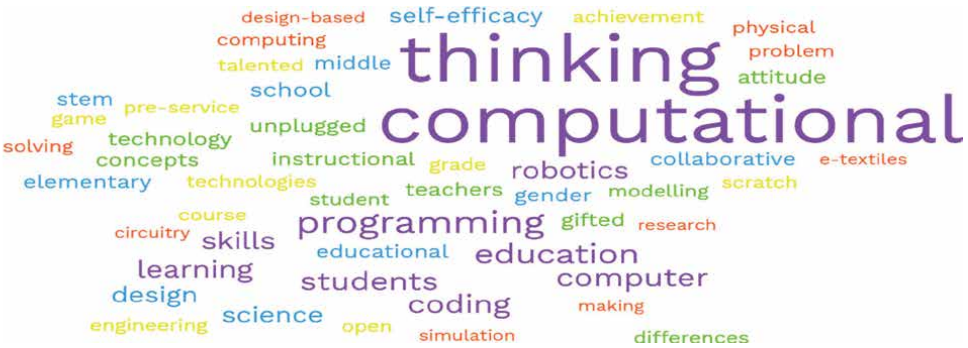


Figure 3.
Keywords related to all the articles under study.

qualitative approach (16), some utilized a quantitative approach (8), and others used a mixed method approach (7). Where studies used some form of intervention to teach CT, a qualitative approach appeared to be appropriate to obtain rich data and understanding of the nuances that emerged. Where quantitative analyses were used, it was generally regarding the perceptions or a scale of the self-efficacies of the main aspects of CT or programming that were needed. The mixed method approach used in some studies allowed for a better understanding of their findings by confirming one set of results with another, especially when design-based or evaluative research was used.

4.3 Interventions and outcomes

To understand the status of CT implementation in teaching, the contexts of interventions and or outcomes in teaching CT in STEM fields were examined. **Table 3** shows the interventions and outcomes of the studies.

While many studies associated CT with programming, other STEM contexts were investigated to develop CT. As can be seen in **Table 4**, almost half of the studies reviewed were conducted in a programming context, while the rest of the studies were conducted in non-programming contexts, but within the STEM fields of specialisation.

What is interesting is that CT was developed using unplugged activities in some studies (2) successfully, which has implications for schools that do not have the computing infrastructure.

Studies that involved students from schools, used a variety of tools and interventions to teach programming with a view to develop CT skills. In most studies this was found to be positive. The following interventions were used: robots, Lego (Wedo2.0, Mindstorms, robotics), Robotics (2), Hopscotch, Scratch, Andruino Scratch, Tinkercard, Educational games (Blocklyscript), and Digital game modding (ChoiCho), as can been seen in **Figure 3**. The outcomes of these studies showed a positive effect on CT development.

In addition to these tools, the studies reviewed indicated the teaching strategies used to develop CT directly in STEM contexts. These interventions are shown on the left of **Figure 3** as independent concepts, namely, Engineering design, Science module, Instructional design, constructing physical model, Web 2.0 tools for digital activities, Unplugged Activities (2) and plugged-in (1), Making projects, Game programming, Design-based learning, Mathematic logic, embodied interaction with technology, Computational modelling, and Tech with Kids web. While it was determined that unplugged activities are able to develop CT [47, 49] effectively, in another study, unplugged activities coupled with plugged-in activities was found to be more effective in improving CT [50].

The claim for introducing CT in compulsory education is based on the notion of transferability of cognitive skills (e.g., logical reasoning and problem-solving abilities). In reviewing the studies, an important aspect in CT advancement is development tools, which concentrates on concepts that assist in understanding CT developmental process, like [20]'s study. Following their approach, the associations between content development tools and concepts, referred to in this article as dependent and independent variables were determined. The findings are presented in **Figure 3**. The main themes (variables) that emerged from the review were computational thinking (CT), programming skills, problem solving, and learning. While some of these studies only considered CT skills, others focussed on the concepts (variables) associated with CT, such as problem-solving, programming, thinking.

SN	Refs.	Interventions	Outcome
1	[26]	Making and maker spaces for making projects supports STEM and developed CT	Positive effect
2	[27]	design of model-building -Throughout the building of the model, students exhibited the use of computational thinking, mainly abstraction, algorithmic thinking, evaluation, generalization, and decomposition.	Positive effect
3	[28]	Findings suggest that this process is a gendered one, with the boys focused more on the operational aspects of building and coding their robots while the girls focused more on group dynamics. Lego robotics	Different effects on girls and boys-using Lego robotics
4	[29]	Organisation of the technology and instructional methods, such as team-based learning, flipped classroom, and pair programming, to help develop CT using Blocked based Hopscotch	Positive effect
5	[30]	Using modifying games with ChoiCo—elements of context-aware integrated CT connecting otherwise fragmented areas such as databases, block-based programming, GIS design.	Positive effect of intervention on CT
6	[31]	The badges, tutorials and some related resources were compiled into the Tech with Kids web to understand CT.	Positive effect
7	[32]	Used computational thinking to build animatronic zoo with coding. Used engineering design as well as coding	CT enabled the design and building
8	[33]	3-week instruction using the science module that had embedded use of CT skills to teach Science.	CT improved the science achievement
9	[34]	game **programming** activities used to scaffold students to learn CT skills. This intervention contributed positively to students' CT skills	game programming has a positive effect on students' CT skills
10	[35]	web 2.0 tools used for digital activities -a significant increase in the participants' **technology** awareness and **computational thinking**	Positive but weak effect on the intended skill. Mainly due to covid

SN	Refs.	Interventions	Outcome
11	[36]	The applications conducted in the Instructional Technologies course and pre-service teachers stated that Scratch applications contributed to the acquisition of Computational Thinking-using Scratch	Significant, positive correlation between CT and academic achievement
12	[37]	learning styles of Web 2.0 based collaborative group activities was used to examine the effects on academic achievement, online cooperative learning attitude level, computer thinking skill level	web 2.0 learning style increased CT
13	[38]	the use of Arduino and Scratch for Arduino (S4A),to study their effect on self-efficacy and motivation towards Science Education, Computational Thinking (CT) and about electricity	Positive effect in view of the conceptual understanding of electricity and CT
14	[39]	To determine the effect of coding instruction performed with the Lego Mindstorms EV3 robotic set on students' attitudes towards coding and their perceptions of computational thinking skills self-efficacy.	positive attitudes towards coding – a significant POSITIVE change in CT skills and self-efficacy perceptions
15	[40]	Attitudes, academic achievements and **computational thinking skills** of the experimental group students, who received robotic-assisted **science education**, toward **science** course differed significantly compared to the students in the control group LEGOWeDo2.0	Positive attitudes, academic achievements, and computational thinking skills with robotic-assisted science education
16	[41]	develop an instructional design that focuses on programming teaching for gifted and talented students and to investigate its effects on the teaching process.	Positive effect the instructional design was effective on CT and creative thinking skills
17	[42]	The open-ended, problem-solving nature of the task requires groups to grapple with the combination of two domains (science and computing) as they collaboratively construct computational models.	CT challenges afford opportunities for students to explore resource-intensive processes, -trial and error, debugging model errors -positive effect
18	[43]	To determine the effect of Tinkercad use in computer programming education on students ' CT skills and perceptions. they were highly motivated for interest and appreciation and found Tinkercad to be generally useful and easy to use	A positive moderate-level relationship between their perception of Tinkercad and their CT skills.

SN	Refs.	Interventions	Outcome
19	[44]	BlocklyScript an EG aims to help students develop their CT by learning basic programming concepts, designing algorithms and correcting mistakes. During the designing phase different EGs were taken under consideration.	The positive results of this pilot evaluation show that BlocklyScript is expected to help students understand CT
20	[45]	the effect of algorithm education on pre-service teachers' computational thinking skills, and computer programming self-efficacy perceptions were examined.10 different algorithmic problems were presented each week, and they were asked to solve these problems using flow chart	There was no effect on CT in general algorithm education had a positive and significant effect only on students' algorithmic thinking
21	[46]	Designed and applied physical computing lessons for elementary 6th-grade students based on the software education guidelines. supported the active interaction of the digital world and the physical world by constructing a physical model using specific media and controlling it with a program.	physical computing lessons materialize students' computational concepts through computational practices, and improve their computational perspectives through the use of authentic contexts
22	[47]	Developed unplugged coding activities using flowcharts for high school students to learn computer science concepts, and to promote their CT skills.	self-directed learning approach used unplugged activities to promote CT
23	[48]	Explored how preservice science teachers used computational thinking as a problem-solving strategy when facilitating IBPW in multiple-deprived classrooms.	positive effect – using CT, they solved problems that otherwise they could not
24	[49]	Examined the role of using unplugged computing activities (based on the Bebras) challenge on developing computational thinking (CT) skills, to promote CT and informatics among school students of all ages.	Students' post-test scores were significantly higher than their pre-test scores
25	[50]	Study aimed to identify the effectiveness of the Chemistry Computational Thinking (CT-CHEM) Module on achievement in chemistry.	Combination of unplugged and plugged-in activities is more effective CT improves achievement in Chemistry
26	[51]	Incorporating a visual execution environment (VEE) and Scratch project for secondary school students as a method to teach and assess computational thinking. Scratch	Knowledge gain on computational and programming concepts and translate CT experiences into reality.

SN	Refs.	Interventions	Outcome
27	[52]	Programming and Debugging—correlation with teacher's rating of executive functions—(EF) is an umbrella term for higher order cognitive functions linked with the frontal lobes of the human brain.	Cognition of CT correlates with programming and debugging activities positively.
28	[53]	Exploring how the CT of two fifth grade learners emerged as an embodied phenomenon during an educational robotics activity.	Robotics activity and embodiment of math concepts, CT emerged
29	[54]	Course in programming a robot for elementary school students and investigated its effectiveness by implementing it in actual classes.	Significantly improved CT thinking and creativity
30	[55]	Examined the correlation between the computer programming self-efficacy and computational thinking skills of students.	Positive effect
31	[56]	The relationship between CT skills and mathematics achievement was statistically significant.	Mathematical logic improves CT skills -positive effect

Table 3.
Interventions and outcomes of the reviewed articles.

Context	Intervention/Tools	Frequency
Programming Context	Operational aspects of building and coding their robots Orchestration of the Technology and instructional methods, such as team-based learning, flipped classroom, and pair programming Lego (WedO2.0, Mindstorms, robotics) Robotics Block programming- Hopscotch, Scratch Educational games (BlockyScript) Creative programming and debugging	15
Non-programming Context	The badges, tutorials and some related resources were compiled into the Tech with Kids web Integrated STEM Mathematical logic Engineering Science Computing Software Chemistry	16

Table 4.
Summary of the context of the studies.

To understand the process of CT skill development, the variables before and after the interventions were considered as well as the relationships between the variables. The dependent and independent variables affecting CT were obtained from the examined studies, separated into themes by content analysis, and then the variable groups and their relationships were determined. The relationships between the variables are shown in **Figure 3**.

The studies addressed dependent variables, such as CT skills, problem solving, and programming skills to measure CT skills. Most of these interventions were based on CT skills and problem-solving variables. However, as shown in **Figure 4**, CT was the dependent variable most frequently studied, followed by programming skills, and problem-solving. Non-programming independent variables, such as the use of digital making and Mathematics learning, games, and Competitive tactile game were mostly used in the studies which included problem-solving skills, as well as CT. The studies that included computer programming as an active independent variable examined the effects on all dependent variables (CT skills, problem solving, and programming skills). Similar, to computer programming, the use of robotics was also associated with most dependent variables. In all studies reviewed, where either robotics or computer programming were used as the intervention tool to promote CT, it was shown to have had a positive effect on CT skills.

In this review, only one study [45], reported no effect on CT skills. This may be related to the duration of the training or intervention, student interest, or the quality of the course. While most studies examined the effect of the intervention on CT,

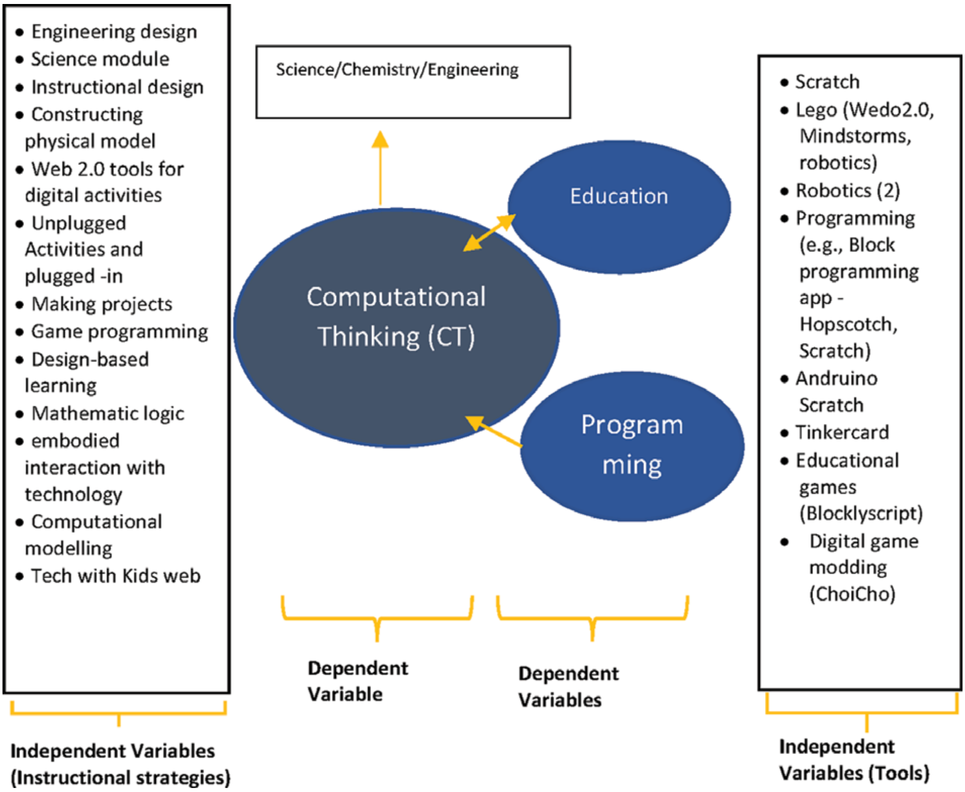


Figure 4.
Relationship of dependent and independent concepts related to CT intervention.

there were three studies [27, 38, 50] that used CT as part of the teaching strategy to determine its effect of students' performance in the related subjects, such as engineering, science, and chemistry. In all three case cases, the outcome was positive.

5. Discussion of results

Several reports have shown the need for the development of computational skills among our current students. They further determined the status of implementing CT in the school curricular. Many studies have shown that coupled with CT, coding or programming or computer science were introduced as compulsory components in the school curriculum at different levels. However, **Table 1**, indicates that while CT is the identifying driver, the context in which it was introduced is the programming or computer science context. From the current review of studies, it was determined that Computational thinking skills can be used in many disciplines, specifically STEM disciplines and is beneficial to all students studying in any field.

Whether it is through an individualized CT course or module, an already existing subject or just as a once-off event, CT can be taught in an enjoyable and engaging way whilst teaching students vital skills which can be applied across the curriculum as well as in daily and work life.

There is no doubt that CT is important and must be developed early in our students. While there are many unplugged activities that can be used, it has been shown that programming is a natural vehicle to develop this skill. In younger learners, use of robotic programming or coding as the buzz word can help to inculcate computational thinking. As has been determined in the reviewed studies, programming or coding and robotics appear to be a major player in the development of CT skills. Since it has been agreed that CT should be developed early in the learner, coding or programming should be taught as compulsory aspects in the curriculum to develop CT skills. However, [30] argue that in practice, it is mostly taught with a narrow focus with just common exercises and testing. An implication is that the teaching strategy needs to change to foreground CT development, with appropriate assessments during programming and coding teaching.

Most if not all sectors of the job market will require some form of coding or programming. Hence, it is important that they can work with algorithmic problem solving and computational methods and tools, a process that should begin in schools. The successful integration of computational thinking concepts into the curriculum requires endeavors in two paths. First, educational policy must be amended to cater to this need and secondly, overcoming infrastructure hurdles, such as the need for teacher resources and teacher education and training. Some emphasis is placed on teacher education regarding the development of CT in STEM related fields [27, 29, 31, 34, 36, 45, 48].

In building problem-solving skills for students, the use of relevant and real problems enhances their understanding, to creatively think of computational steps towards a solution. Hence, it is important to design a learning tool that allow users to teach/learn programming concepts through CT approach while abstracting problems that are familiar within a context.

5.1 Limitations

Some limitations need to be acknowledged. Firstly, a small number of publications (31) qualified for inclusion, and the database search was restricted to just the ERIC database as explained earlier. Hence, other databases of repute should be used as well and more research needs to be carried out to confirm or otherwise, the findings. A potential bias for the study is the influence the researcher had upon the analysis, despite screening the articles at least three times for the final review. Although the current study is based on a limited sample, this work offers valuable insights into the status of CT in STEM and lessons in developing CT in our students.

6. Conclusion

The study set out to determine the trends and status of computational thinking in STEM fields, by problematizing the lack of development strategies of CT within STEM fields. A systematic review was followed. What has become clear is awareness of the need for increasing the uptake of STEM subjects and the acknowledgement of computational skill as a twenty-first century skill. It has been established that CT is a necessary skill to develop in previous studies and in the current view. Several interventions to teach and develop CT in the STEM fields have been explored, indicating robotics as a driver for primary school children to learn CT. It has become abundantly clear that programming and coding with robotics appear to be most used for the development of CT are key to fostering of CT skills. To conclude, the results of this study indicate that there is much work to do regarding teacher education to promote CT skills in their curricular. Despite the emerging research specifically in CT in STEM fields, more needs to be done. Research in CT disciplinary pedagogical studies and transdisciplinary studies need to be conducted.

Author details

Irene Govender
University of KwaZulu-Natal, Durban, South Africa

*Address all correspondence to: govenderi4@ukzn.ac.za

References

[1] Li Y, Schoenfeld AH, diSessa AA, Graesser AC, Benson LC, English LD, et al. On computational thinking and STEM education. Journal of STEM Education Research. 2020;**3**:147-166. DOI: 10.1007/s41979-020-00044-w

[2] Karpudewan M, Law KE, Zaharudin R. Computational thinking in STEM education among matriculation science students. Asia Pacific Journal of Educators and Education. 2021;**36**(1): 177-194. DOI: 10.21315/apjee2021.36.1.10

[3] Caeli EN, Yadav A. Unplugged approaches to computational thinking: A historical perspective. TechTrends. 2020;**64**:29-36. DOI: 10.1007/s11528-019-00410-5

[4] Bocconi S, Chioccariello A, Dettori G, Ferrari A, Engelhardt K. Developing computational thinking in compulsory education – Implications for policy and practice. EUR. 2016:28295. DOI: 10.2791/792158

[5] Papert S. Teaching children thinking. Contemporary Issues in Technology and Teacher Education. 1971;**5**(3/4):353-365

[6] Li Y, Wang K, Xiao Y, Froyd JE. Research and trends in STEM education: A systematic review of journal publications. International Journal of STEM Education. 2020;**7**:11. DOI: 10.1186/s40594-020-00207-6

[7] Wing JM. Computational thinking's influence on research and education for all. Italian Journal of Educational Technology. 2017;**25**(2):7-14. DOI: 10.17471/2499-4324/922

[8] Sylo MM. From Algorithmic to Computational Thinking: On the Way for Computing for all Students. ITICSE'15. Vilnius, Lithuania: ACM; 2015. DOI: 10.1145/2729094.2742582

[9] Hermans F, Aivaloglou E. To Scratch or not to Scratch? A controlled experiment comparing plugged first and unplugged first programming lessons. In: Proceedings of the 12th Workshop in Primary and Secondary Computing Education. 2017. pp. 49-56

[10] Wing JM. Computational thinking. Communication of the ACM. 2006;**49**(3):33-35. DOI: 10.1145/1118178.1118215

[11] Voogt J, Fisser P, Good J, Mishra P, Yada A. Computational thinking in compulsory education: Towards an agenda for research and practice. Educational Information Technology. 2015;**20**:715-728. DOI: 10.1007/s10639-015-9412-6

[12] Yadav A, Stephenson C, Hong H. Computational thinking for teacher education. Communications of the ACM. 2017;**80**(4):55-62

[13] Barth-Cohen LA, Jiang S, Shen J, Chen G, Eltoukhy M. Interpreting and navigating multiple representations for computational thinking in a robotics programing environment. Journal of STEM Education Research. 2018;**1**(1):119-147

[14] Hsu T-C, Chang S-C, Hung Y-T. How to learn and how to teach computational thinking: Suggestions based on a review of the literature. Computers & Education. 2018;**126**:296-310

[15] Shute VJ, Sun C, Asbell-Clarke J. Demystifying computational thinking. Educational Research. 2017;**22**:142-158

[16] Bocconi S, Chioccariello A, Earp J. The Nordic approach to introducing Computational Thinking and programming in compulsory education. Report prepared for the Nordic@BETT

2018 Steering Group. 2018. pp. 1-32. DOI: 10.17471/54007

[17] Heintz F. From computational thinking to AI. TedX Talks. 2017; YouTube

[18] Yadav A, Mayfield C, Zhou N, Hambrusch S, Korb JT. Computational thinking in elementary and secondary teacher education. ACM Transactions on Computing Education (TOCE). 2014;**14**(1):5

[19] Alegre F, Underwoood J, Moreno J, Alegre M. Introduction to Computational Thinking: A New High School Curriculum using CodeWorld. SIGCSE '20, March 11-14. 2020; Portland, OR, USA. 10.1145/3328778.3366960

[20] Taslibeyaz E, Kursun E, Karaman S. How to develop computational thinking: A systematic review of empirical studies. Informatics in Education. 2020;**19**(4):701-719. DOI: 10.15388/infedu.2020.30

[21] Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med. 2009;**6**:e1000097

[22] Li Y, Froyd JE, Wang K. Learning about research and readership development in STEM education: A systematic analysis of the journal's publications from 2014 to 2018. International Journal of STEM Education. 2019;**6**(19):1-8. DOI: 10.1186/s40594-019-0176-1

[23] Xu M, Williams PJ, Gu J, Zhang H. Hotspots and trends of technology education. International Journal of Technology and Design Education. 2020;**30**:207-224. DOI: 10.1007/s10798-019-09508-6

[24] Brown J. The current status of STEM education research. Journal of STEM Education: Innovations & Research. 2012;**13**(5):7-11

[25] Mizell S, Brown S. The current status of STEM education research 2013-2015. Journal of STEM Education: Innovations & Research. 2016;**17**(4):52-56

[26] Hébert C, Jenson J. Making In Schools: Student Learning Through An E-Textiles Curriculum. Discourse: Studies in the Cultural Politics of Education. 2020;**41**(5):740-761

[27] Lyon JA, Magana AJ. The use of engineering model-building activities to elicit computational thinking: A design-based research study. Journal of Engineering Education. 2021;**110**(1):184-206

[28] Ardito G, Czerkawski B, Scollins L. Learning computational thinking together: Effects of gender differences in collaborative middle school robotics program. TechTrends: Linking Research and Practice to Improve Learning. 2020;**64**(3):373-387

[29] Zha S, Jin Y, Moore P. Hopscotch into coding: Introducing pre-service teachers computational thinking. TechTrends: Linking Research and Practice to Improve Learning. 2020;**64**(1):17-28

[30] Kynigos C, Grizioti M. Modifying games with ChoiCo: Integrated affordances and engineered bugs for computational thinking. British Journal of Educational Technology. 2020;**51**(6):2252-2267

[31] Hunsaker E, West RE. Designing computational thinking and coding for early childhood educators. TechTrends: Linking Research and Practice to Improve Learning. 2020;**64**(1):7-16

[32] Deniz H, Kaya E, Yesilyurt E. Integrating engineering, science,

reading, and robotics across grades 3-8 in a STEM education era. Journal of Learning and Teaching in Digital Age. 2021;**6**(1):40-45

[33] Lapawi N, Husnin H. The effect of computational thinking module on achievement in science. Science Education International. 2020;**31**(2):164

[34] Kukul V, Çakir R. Exploring the development of primary school students' computational thinking and 21st century skills through scaffolding: Voices from the stakeholders. International Journal of Computer Science Education in Schools. 2020;**4**(2):1-17

[35] Çevik M, Baris N, Sirin M. The effect of digital activities on the technology awareness and computational thinking skills of gifted students (eTwinning project example). International Journal of Modern Education Studies. 2021;**5**(1):205-244

[36] Ilic U. The impact of scratch-assisted instruction on computational thinking (CT) skills of pre-service teachers. International Journal of Research in Education and Science. 2021;**7**(2):426-444

[37] Pürbudak A, Usta E. Collaborative group activities in the context of learning styles on web 2.0 environments: An experimental study. Participatory Educational Research. 2021;**8**(2):407-420

[38] Ntourou V, Kalogiannakis M, Psycharis S. A study of the impact of arduino and visual programming in self-efficacy, motivation, computational thinking and 5th grade students' perceptions on electricity. EURASIA Journal of Mathematics, Science and Technology Education. 2021;**17**(5):1-11

[39] Yildiz T, Seferoglu SS. The effect of robotic programming on coding attitude and computational thinking skills toward self-efficacy perception. Journal of Learning and Teaching in Digital Age. 2021;**6**(2):101-116

[40] Usengül L, Bahçeci F. The effect of LEGO WeDo 2.0 education on academic achievement and attitudes and computational thinking skills of learners toward science. World Journal of Education. 2020;**10**(4):83-93

[41] Avcu YE, Er KO. Developing an instructional design for the field of ICT and software for gifted and talented students. International Journal of Educational Methodology. 2020;**6**(1):161-183

[42] Emara M, Hutchins NM, Grover S. Examining student regulation of collaborative, computational, problem-solving processes in open-ended learning environments. Journal of Learning Analytics. 2021;**8**(1):49-74

[43] Eryilmaz S, Deniz G. Effect of tinkercad on students' computational thinking skills and perceptions: A case of Ankara province. Turkish Online Journal of Educational Technology – TOJET. 2021;**20**(1):25-38

[44] Karakasis C, Xinogalos S. BlocklyScript: Design and pilot evaluation of an RPG platform game for cultivating computational thinking skills to young students. Informatics in Education. 2020;**19**(4):641-668

[45] Türker PM, Pala FK. The effect of algorithm education on students' computer programming self-efficacy perceptions and computational thinking skills. International Journal of Computer Science Education in Schools. 2020;**3**(3):1-11

[46] Min SH, Kim MK. Developing children's computational thinking through physical computing lessons. International

Electronic Journal of Elementary Education. 2020;**13**(2):183-198

[47] Threekunprapam A, Yasri P. Patterns of computational thinking development while solving unplugged coding activities coupled with the 3s approach for self-directed learning. European Journal of Educational Research. 2020;**9**(3):1025-1045

[48] Tsakeni M. preservice teachers' use of computational thinking to facilitate inquiry-based practical work in multiple-deprived classrooms. EURASIA Journal of Mathematics, Science and Technology Educatio. 2021;**17**(1):em1933

[49] Delal H, Oner D. Developing middle school students' computational thinking skills using unplugged computing activities. Informatics in Education. 2020;**19**(1):1-13

[50] Chongo S, Osman K, Nayan NA. Impact of the plugged-in and unplugged chemistry computational thinking modules on achievement in chemistry. EURASIA Journal of Mathematics, Science and Technology Education. 2021;**17**(4):em1953

[51] Hijón NR, García-Iruela M, Connolly C. Developing and assessing computational thinking in secondary education using a TPACK guided scratch visual execution environment. International Journal of Computer Science Education in Schools. 2021;**4**(4):1-16

[52] Robertson J, Gray S, Toye M, Booth JN. The Relationship between Executive Functions and Computational Thinking. International Journal of Computer Science Education in Schools. 2020;**3**(4):1-15

[53] Kopcha TJ, Ocak C, Qian Y. Analyzing children's computational thinking through embodied interaction with technology: A multimodal perspective. Educational Technology Research and Development. 2021;**69**(4):1987-2012

[54] Noh J, Lee J. Effects of robotics programming on the computational thinking and creativity of elementary school students. Educational Technology Research and Development. 2020;**68**(1):463-484

[55] Avcu YE, Ayverdi L. Examination of the computer programming self-efficacy's prediction towards the computational thinking skills of the gifted and talented students. International Journal of Educational Methodology. 2020;**6**(2):259-270

[56] Chongo S, Osman K, Nayan NA. Level of computational thinking skills among secondary science student: Variation across gender and mathematics achievement. Science Education International. 2020;**31**(2):159-163

Chapter 2

Impact of Dialogic Argumentation Pedagogy on Grade 8 Students' Epistemic Knowledge of Science

Getachew Tarekegn, Jonathan Osborne and Mesfin Tadesse

Abstract

This study explores the effect of dialogic argumentation on grade 8 students' epistemic knowledge of science in physics. A quasi-experimental design was employed to compare experimental (239) and control (240) groups' epistemic knowledge of science. A pre-intervention and post-intervention physics reasoning test was administered, and small group classroom discussions were also video recorded. Physics teachers in the intervention group had trained for three days about dialogic argumentation and *Talking Physics Students Activities manual* was also distributed and used in this yearlong intervention. Mann-Whitney U test results indicated that the post-test scores of grade 8 students in the argumentation lessons significantly increased in their level of epistemic knowledge compared to the non-argumentation groups, $z = -4.509$, $p = .000$, and $r = .21$, but not in the pre-test scores, $z = -1.038$ and $p = .299$. However, both pre- and post-test scores of both groups were relatively low. The intervention groups showed significant improvements in the quality of their argumentation on the ASAC scale, $z = 2.111$, $p = .035$, and $r = .56$, but not the control groups, $z = 1.068$ and $p = .285$. The study found evidence that argumentation-based lessons improved both the epistemic knowledge and the quality of dialogic argumentations of grade 8 students and that students' level of epistemic knowledge and the quality of their dialogic argumentations were strongly correlated.

Keywords: dialogical, argumentation, physics education, epistemic knowledge, scientific reasoning

1. Introduction

Research in science education indicates that dialogical argumentation is becoming more popular as a beneficial instructional approach for motivating and promoting students' reasoning skills and their ability to rebuild basic scientific concepts. Dialogical argumentation is a type of classroom debate in which two or more arguers hold opposing viewpoints on a given topic. The goal is to use evidence to support or disprove claims, to socially build concepts through language mediation, and to improve collective understanding [1]. Such discourse, it may be argued, helps students improve their competencies to "explain phenomena scientifically, evaluate and design scientific inquiry, and interpret data and evidence scientifically" [2]. These competencies, in turn, necessitate the integration of three types of scientific knowledge. The first is *content knowledge*, which is a set of domain-specific concepts and theories that serve as the foundation for scientific reasoning. The second type of scientific knowledge, referred to as *procedural knowledge*, is an understanding of the procedures and techniques used to back up what science claims to know. *Epistemic knowledge*, the third type, is knowledge of epistemic constructs, as well as the validation and justification of the claims of science [2–5]. Many studies have attested the success of dialogical argumentation in improving conceptual knowledge in Africa and elsewhere [6–10]. However, there are few studies in science education that focus on the impact of dialogical argumentation on epistemic knowledge of science.

This study looks into how students' epistemic knowledge of science is transformed through dialogical argumentation. As reform documents have repeatedly stated, epistemic knowledge of science is a critical factor in introducing science-as-a-practice in science education [2]. However, the development of scientific knowledge in science education has traditionally focused on content knowledge rather than epistemic and procedural knowledge. On the other hand, the goals of learning science are shifting away from outcomes, such as conceptual learning, problem solving, and science process skills, and toward introducing and developing the competencies and tools needed for scientific knowledge generation and construction [11]. This shift has highlighted the importance of evidence-based argumentation and the role of epistemic knowledge in science.

Developing epistemic knowledge is especially crucial in the cultural context of Africa, where knowledge is commonly constructed on the basis of authoritative statements, such as those made by teachers, elders, or books. Traditional culture in Ethiopia, where this study was conducted, promotes a stereotyped and authoritative view of knowledge [12]. As a result, science teachers perceive and describe scientific knowledge as everlasting truth, a viewpoint that obstructs teachers' comprehension of how science works and, as a result, the teaching and learning of school science. The dominating teacher-centered approach to science teaching, the prevalence of science classrooms devoid of scientific inquiry, and rote-memory-based school science tests are all-natural extensions of this culture. Africa needs to make significant efforts to transform the nature of its science education to profoundly alter both instructors' and students' innate beliefs about knowledge [7]. To effect a paradigm shift in science pedagogy, teachers' and students' epistemic knowledge must be significantly improved. In conclusion, past research shied away from a thorough examination of the progression of the three parts of scientific knowledge, notably epistemic knowledge of science. These studies have failed to explore the impact of dialogical argumentation on the development of epistemic knowledge. In this study, we look at how middle school students' epistemic knowledge is influenced by dialogical argumentation.

The purpose of this study is, thus, to examine the effect of middle school students' engagement in dialogical argumentation on their epistemic knowledge of science. Specifically, the research questions guiding this study are:

1. To what extent does engagement in dialogical argumentation affect middle school students' epistemic knowledge?

2. Is there a relationship between the quality of dialogical argumentation generated by middle school students and their level of epistemic knowledge?

This study was part of a wider research project, namely Transforming Pedagogy of STEM Subjects (TPSS), which promotes the transformation of STEM pedagogy in middle schools of Ethiopia through dialogical argumentation.

2. Epistemic knowledge and dialogical argumentation

In this section, we review the existing body of knowledge that focuses on dialogical argumentation and epistemic knowledge.

2.1 Epistemic knowledge

Epistemic knowledge is the knowledge of how we know what we know [2, 4]. Epistemic knowledge is essential to constructing and justifying the knowledge produced by science. Such knowledge empowers one's functional understanding of the nature of science [13, 14] and its implementation introduces students to the core scientific practices [5]. Knowing and identifying scientific constructs, such as hypothesis and theory, as well as justifying ideas with evidence and comprehending the justification for various modes of scientific research, are all instances of epistemic knowledge [15, 16].

There are several distinct and changeable parameters in the natural world, which are referred to as variables. Variables are identified and isolated to understand their distinct functions and contributions to the observable world. To understand what a given variable contributes to the observed phenomena, it is necessary to control other factors and examine the change brought about by that specific variable. The generation of knowledge in science necessitates an understanding of why a specific method is used, as well as how the method is both systematic and repeatable. This type of understanding is part of epistemic knowledge [4] and students armed with knowledge can identify entities and events in the physical world, create explanatory models of causal links, and explain how they know what they know.

Many academics see science as a body of knowledge and a way of knowing that takes time to develop [13, 17]. Before science's methodologies acquired the current systemic ways of knowing, it had to travel a long and difficult road. Indeed, many stories of scientists' faulty viewpoints and incorrect explanations and reasoning advanced to explain many natural phenomena can be found throughout the history of science [17]. Justifications for knowledge claims rely primarily on data-based evidence and the quality of the reasoning employed to support scientific arguments. Depending on the focus of investigations and the nature of the data, such arguments may be hypothetico-deductive (e.g., Copernicus' argument for the heliocentric system), inductive (e.g., the conservation of energy), or inference to the best explanations (e.g., Darwin's theory of evolution or Wegener's argument for moving continents) [2].

Insights into the role and nature of such arguments are critical for the development of scientific literacy, which needs students to understand the constructs and distinguishing qualities of science [2]. Understanding the underpinnings of scientific observations, facts, hypotheses, models and theories, purposes, goals and values of science, and the styles of reasoning used are integral to grasping science. Thus, epistemic knowledge enables us to understand the role and goal of inquiry in producing knowledge, the goal of the inquiry, and the methodology of the inquiry to justify the claims to know. We would, therefore, argue that understanding science requires understanding how we think in science.

2.2 Characterizing features of epistemic knowledge

As previously stated, epistemic knowledge is essential to the scientific process of knowledge development. Epistemic knowledge provides students with an understanding of the constructs employed in scientific reasoning, as well as the constituent aspects of these constructs and their significance in explaining scientific knowledge. Students can appreciate the intellectual achievement of scientific ideas and their growth using this knowledge. Another benefit of epistemic knowledge is that it allows students to think more deeply about the natural phenomenon they are studying through the use of constructs, such as hypothesis, theory, facts, models, and theories. These constructs are important for developing model-based explanations of how the natural world operates, which is what science is all about. Epistemic knowledge also assists students in understanding what constitutes a scientific question and what observations may yield relevant data. Here, the student's epistemic knowledge is crucial in selecting the appropriate method of empirical investigation and explaining the resulting design. After that, students must analyze and interpret data to determine how the data informs their hypotheses and theories. Students also must ask questions, such as: Do they support one idea over others, help to refute an idea, or suggest an entirely new explanation? In short, epistemic knowledge is the knowledge that explains how scientific claims are supported by data and reasoning in science. Furthermore, epistemic knowledge includes an understanding of how measurement uncertainty affects our level of confidence in our claims to know. Students must also apply several styles of reasoning in science, including deductive, inductive, abductive, analogical, and model-based reasoning. Understanding their role in science requires epistemic knowledge [2–5, 16, 18, 19].

3. Dialogical argumentation: what is it?

Dialogical argumentation is a discourse that involves reasoning to solve a problem collaboratively and resolve a conflict between ideas [20]. Dialogical argumentation uses dialogs to justify or refute claims based on evidence, construct ideas socially through the mediation of language, and enhance shared understanding [21]. In the classroom, students can engage in argumentative discourses either at an individual or small group level. During the whole class discussion, individual or group leaders' views can be presented to the group or the whole class, and group leaders or teachers reflect and mediate the co-construction of knowledge.

Dialogical argumentation provides an opportunity for students to see that a problem or a concept may be comprehended in various ways. Each contender of each competing view forwards its claims and supports these claims with evidence or reason. Not convinced arguers can make counterclaims and justify it with reasons or make rebuttals. Through this approach, dialogical argumentation supports students' scientific literacy, conceptual learning, and improves their epistemic understanding. This section presents some justifications to support dialogical argumentation in science classrooms.

Several studies thus far have linked argumentation with scientific literacy [22–25]. Many curricula reforms encourage the incorporation of scientific issues relevant to the well-being of society [26, 27]. Using dialogical argumentation, students can evaluate socio-scientific issues and reach evidence-based conclusions. This, in turn, contextualizes, humanizes, and socializes science for students [17]. Involving students in dialogical argumentation about socio-scientific issues have many benefits. First, students learn to structure their arguments in the context of socio-scientific issues. Second, students know how to justify their stances using normative scientific criteria. Third, students learn to take well throughout evaluative judgments.

Dialogical argumentation situates talking at the center of classroom discourse. Provided that dialogical argumentation is dialogic in nature, it also aligns with the socio-cultural view of learning [28, 29]. Alexander [30] mentioned that effective learning requires the interconnection between language, thinking, and knowing. Therefore, classroom discourses should focus on strategies that promote students' reasoning. Dialogical argumentation is one of such teaching approaches that connect these three constructs mentioned earlier together [31]. Moreover, dialogical argumentation challenge what students know. When students reflect their knowledge, cognitive conflict, which is key to conceptual learning, might surface.

Studies in the history and philosophy of science revealed that scientists often negotiate meaning in their attempt to understand a scientific event. In addition, scientists critically consider and discuss alternative conceptions and competing views to convince their claims to the scientific community. In this manner, Kim and Roth [28] put argumentation at the heart of the practice of science. Kelly and Licona [32] further argue for school science curricula that value and address the various scientific practices discussed in the history and philosophy of science studies, such as modeling, inquiry, and argumentation that played a paramount role in advancing science. These practices should then be represented in school science curricula, with an emphasis on involving students in scientific activities [33]. Argumentation has thus emerged as a key productive strategy in making sense of physics phenomena and partaking in epistemic practices of science.

4. Promoting epistemic knowledge with dialogical argumentation

This study proposes that students' engagement in dialogical argumentation enhances students' epistemic knowledge of science. As discussed elsewhere, during dialogical argumentation, students experience the value to consider alternative positions, justifying their own views with evidence to persuade others, and challenging their arguers' positions [34]. These experiences create favorable conditions for students to facilitate their epistemic understanding. Engaging in considering alternative views, which are supported by evidence, could induce students to think of the idea that there may be more than a single objective reality. In addition, when students give and receive a critique of their ideas, students could recognize the value of argumentation and evidence in constructing scientific knowledge. Previous research has indicated that students' engagement in dialogical argumentation improves students' epistemic practices of science, such as providing justifications and justifying claims with evidence and reason [35, 36].

Particularly the findings of Iordanou and Constantinou [35] indicated that as a result of dialogical argumentation, students use and refer to data to support their arguments. In addition, these authors reported a decrease in personal opinions during the argumentative discourse. A shift from unsupported claims toward using evidence to support their claims and providing interpretations for other evidence was observed. The subtle implication of these findings is that argumentation promotes epistemic practices of students. In another study, Kuhn, Zillmer [36] Kuhn et al. confirms that students show improvements in their epistemic understanding of the normative nature of arguments after engaging in dialogical argumentation. The gain in epistemic understanding includes an understanding of what are acceptable claims to knowledge and what are acceptable ways of discourse in science. However, these studies report the progress of epistemic understanding of argumentation without directly measuring epistemic understanding.

Erduran and Dagher [13] claim that including students in argumentation allows them to see science as a logical practice and to assess the strengths and shortcomings that occurred during the process of knowledge construction in science. In short, such rationales position argumentation as an epistemic activity and emphasize the importance of incorporating argumentation as one of the foci of science teaching. The epistemic process involved in developing and evaluating scientific knowledge claims is revealed via argumentation. This epistemic process, in turn, employs epistemic criteria to select evidence, assess claims and evidence, produce counter-claims, and effectively communicate with others. Hence, it is important to assess the epistemic knowledge of students and to investigate the influence of engaging students in dialogical argumentation on their epistemic practices of science. Thus, this study will, therefore, explore how middle school students' engagement in dialogical argumentation influences their epistemic knowledge of physics.

5. Methodology

5.1 About participants

A convenience sampling technique was used to identify the primary schools. Twelve primary schools, which use English as a medium of instruction, accessible to transport and cooperative to participate in this yearlong study, were identified and participated. All these schools are government schools. In total, 14 classrooms were randomly selected from these schools and assigned as experimental (N = 7) and control (N = 7) groups. From these schools, a total number of 479 Grade 8 students were involved in this study. Among these 479 middle school students, 273 (57.0%) were female and 206 (43.0%) were male. The treatment group comprised 239 (49.9%) students, whereas the control group contained 240 (50.1%) Grade 8 students.

5.2 Research design: mixed methods

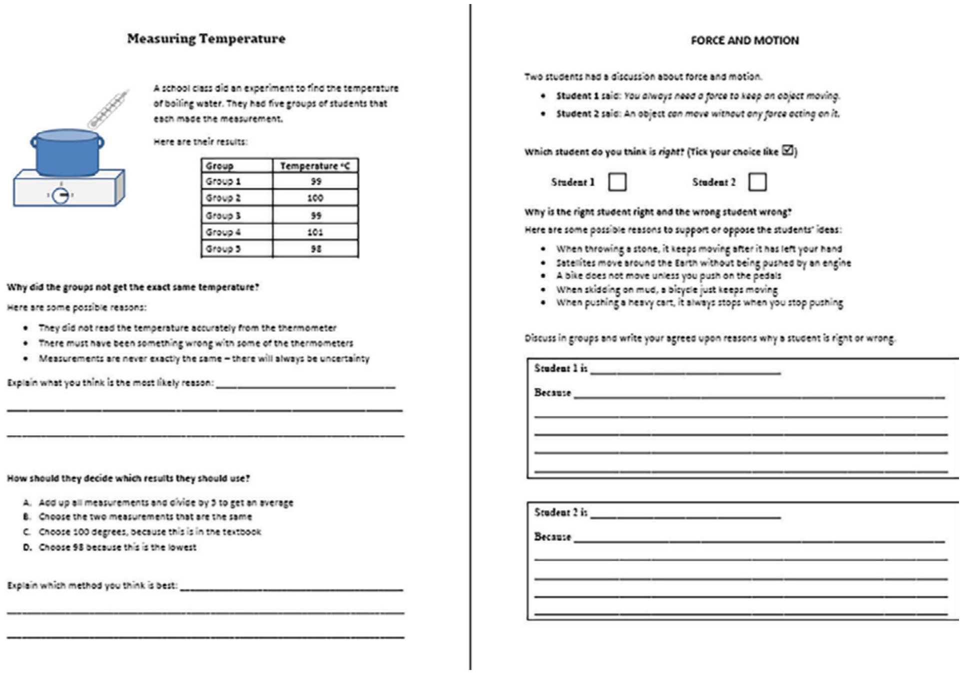
We employed a quasi-experimental mixed-method design that combines quantitative postpositivist methods with qualitative constructivist methods. Study subjects, eighth-grade students, were split into an argumentation lesson group and a non-argumentation lesson group, and quantitative data about students' level of epistemic knowledge were collected from both groups before and after intervention using Epistemic Knowledge Test items. Students' argumentations were video recorded to gather qualitative data. According to Creswell and Clark [37], a mixed-methods experimental (or intervention) design is an "approach in which the researcher embeds the collection, analysis, and integration of both quantitative and qualitative data within an experimental quantitative research design" (p. 139). To better analyze and interpret the effect of dialogic argumentation on students' epistemic knowledge, this study needs not only quantitative data from epistemic knowledge test scores but also qualitative data from observations of students' argumentation. The use of quantitative and qualitative methods together increases the validity of research reports by correlating results from different methods, elaborating results from one method, informing the other method, providing a comprehensive account of results reported by one method, and even contradicting results reported by one method [38]. These features made mixed-methods relevant and appropriate for this study.

Furthermore, the mixed methods experimental (or intervention) design was chosen since the goal of the study was to learn not only about the impact of argumentation classes on epistemic knowledge but also about how the intervention works. we adopted a concurrent embedded mixed-methods experimental design in which quantitative and qualitative data are collected at the same time. To put it briefly, the study adopted a mixed-method experimental design by integrating quantitative and qualitative approaches to science education research [39].

5.3 Context of the study

This study was part of a wider research project called Transforming Pedagogy of STEM Subjects (TPSS), which promotes the transformation of STEM pedagogy in Ethiopia through argumentation. The aim of TPSS was to reduce an overreliance on teacher-centered, didactic teaching and increase the use of students-oriented, dialogical teaching.

TPSS deduced two strategies to achieve its goal. The first strategy was implementing dialogical teaching and the second strategy was implementing dialogical argumentation directly into elementary schools through in-service training. The participant of this study was part of the group addressed in the second strategy. In 2017/2018, TPSS runs a training program for Grade 7 and 8 physics teachers to



Measuring Temperature

A school class did an experiment to find the temperature of boiling water. They had five groups of students that each made the measurement.

Here are their results:

Group	Temperature °C
Group 1	99
Group 2	100
Group 3	99
Group 4	101
Group 5	98

Why did the groups not get the exact same temperature?

Here are some possible reasons:

- They did not read the temperature accurately from the thermometer
- There must have been something wrong with some of the thermometers
- Measurements are never exactly the same – there will always be uncertainty

Explain what you think is the most likely reason: ____________________

____

____

How should they decide which results they should use?

A. Add up all measurements and divide by 5 to get an average
B. Choose the two measurements that are the same
C. Choose 100 degrees, because this is in the textbook
D. Choose 98 because this is the lowest

Explain which method you think is best: ____________________

____

____

FORCE AND MOTION

Two students had a discussion about force and motion.

- **Student 1** said: *You always need a force to keep an object moving.*
- **Student 2** said: An object *can move without any force acting on it.*

Which student do you think is *right*? (Tick your choice like ☑)

Student 1 ☐ **Student 2** ☐

Why is the right student right and the wrong student wrong?

Here are some possible reasons to support or oppose the students' ideas:

- When throwing a stone, it keeps moving after it has left your hand
- Satellites move around the Earth without being pushed by an engine
- A bike does not move unless you push on the pedals
- When skidding on mud, a bicycle just keeps moving
- When pushing a heavy cart, it always stops when you stop pushing

Discuss in groups and write your agreed upon reasons why a student is right or wrong.

Student 1 is ____________________

Because ____________________

____

____

____

____

Student 2 is ____________________

Because ____________________

____

____

____

____

Figure 1.
Scaffolding activities used for video-recorded students' discourse: Activity for 1st (measuring temperature) and 2nd videoing (force and motion).

incorporate argumentation and a more dialogic approach to their physics teaching. In the three-day training, elementary physics teachers analyzed physics teaching, student learning, and the physics curriculum of Ethiopia. Then, they were introduced to scientific reasoning and argumentation theory. They had been also trained on how to stimulate argumentation in their classroom. TPSS uses scaffolding activities, similar to **Figure 1**, which make students reflect and discuss scientific ideas. This motivates students to adopt specific ways of talking and expressing scientific ideas, and participate in a structured cultural discourse, that is, scientific argumentation, which promotes their learning. The scaffolding activities address common misconceptions identified via science education research and depend on a variety of styles of scientific reasoning typically used in scientific argumentation.

6. Overview of the grade 7 and 8 physics curriculum

Physics is one of the natural science disciplines taught in Ethiopian schools in Grades 7 and 8. It was intended that learning physics would help students understand the physical world, conduct observations and experiments relating to physical events and phenomena, and increase their interest in the natural world. Furthermore, the curriculum highlights the significance of physics in the study of other STEM topics as well as students' comprehension of scientific practices. The student-centered teaching method aims to help students acquire scientific knowledge, skills, and attitudes and build their confidence to apply knowledge gained to solve text-book or real-life problems. To sum up, the Federal Democratic Republic of Ethiopia Ministry of Education has mentioned the following advantages of studying physics at the K-12 levels. These include:

- Understanding the functioning principles of many of our everyday utensils and gadgets'
- learning about nature and how it works,
- applying physics expertise to other fields and disciplines, and
- addressing practical challenges in the real world [26].

So, to understand nature and natural phenomena, physics should be taught by engaging students in scientific practices rather than telling facts to students. The school science curriculum framework clearly indicated that it is the learner-centered approaches and the constructivist epistemology that the teaching-learning process should follow [26, 40]. The Ministry of Education stresses the importance of quality teaching and learning materials to bring quality to science and mathematics education. Students' textbooks, students' workbooks, teacher's guides, teacher's handbooks, syllabi, minimum learning competency guides, audio-video materials, and other teaching-learning resources are taken to be important in improving science and mathematics education.

Topics covered in the Grade 7 physics textbooks in Ethiopian schools are structured in eight units: Physics and Measurement, Motion, Force and Newton's Laws of Motion, Work, Energy and Power, Simple Machines, Temperature & Heat, Sound, and Electricity & Magnetism. Grade 8 physics textbook is organized into six units and covers the topics: Physics and Measurement, Motion in One Dimension, Pressure, Heat Energy, Electricity & Magnetism, and Light.

7. About items that measure epistemic knowledge

The physics reasoning test was developed based on a theoretical rationale presented by Kind and Osborne [4]. The test includes items for different styles of reasoning—hypothetical modeling, experimenting, and mathematical-deductive reasoning. In addition, the physics reasoning test is more focused on items allocated to each of the three sub-constructs of scientific knowledge, namely—content, procedural, and epistemic knowledge. A physics reasoning test that contains 20 questions were administered to each group at the beginning and end of the intervention. The physics reasoning test was prepared to measure students' scientific knowledge, which includes conceptual, procedural, and epistemic knowledge. Since the focus of this study is on epistemic knowledge, we identified items that assess the epistemic knowledge of students.

Among physics reasoning test items, 10 items had been identified as an item that measures epistemic entity of scientific knowledge based on features identified by [2, 4, 5]. These are Item 3, Item 4, Item 5, Item 6, Item 7, Item 8, Item 9, Item 10b, Item 11, and Item 12 (see the Appendix). To identify items that assess epistemic knowledge, four PhD candidates in Science Education at Addis Ababa University were employed. The PhD candidates took courses, namely Scientific Reasoning and Argumentation in Science Education and Assessment in Physics Education; hence, we considered them appropriate to identify items that measure epistemic knowledge. We asked them to categorize each of the 20 items of the physics reasoning test based on the kind of scientific knowledge (either content, procedural, epistemic knowledge, or the combination of any of these). The raters evaluated each of the test items and identified the knowledge the test item was supposed to measure. We compared the scores generated to measure the inter-rater reliability of the categorization of the physics reasoning test with Cohen's kappa, which was .833. This indicates that it is good to have an excellent agreement between the identification of each item with the knowledge type it assesses.

We automatically assigned an item to its specific knowledge group when an item is categorized in that group by three or more raters. When raters were split about a particular item, discussions were held to clarify and decide using a panel of experts discussion. Twenty-three items were categorized similarly by three or more raters while the remaining three items were split. Further discussions were conducted and two of the items were assigned in the epistemic knowledge category, while the remaining were categorized as procedural knowledge. To sum up, the physics-reasoning test contains three, 12, and 10 items that measure the content, procedural, and epistemic knowledge, respectively. In this study, we used only these 10 epistemic knowledge items (EKIs hereafter). The maximum possible score of the 10 EKIs was 15 points. We present items of EKIs in the Appendix.

The chosen EKIs examine the epistemic construct of deduction as a form of argument in the construction of scientific knowledge, the knowledge required for evaluating scientific experimentation, and the role of scientific theories and ideas in model-based scientific reasoning. The EKIs items are designed to assess how students know what they know in science. All of the items are applicable to Ethiopia's physics curriculum for middle school students. To capture the notion of epistemic knowledge, we followed the conception of epistemic knowledge as outlined in Refs. [2, 4, 5]. Thus, items in EKIs assess students' reasoning about how they know that a certain physics concept and phenomenon is right or wrong, their understanding of how data is represented and used to scientifically justify claims and conclusions using evidence, and their knowledge about making inferences.

8. The intervention

The intervention in the middle school began with a visit to local education officers and school directors, where they were informed of the TPSS aims and research plans. We gained authorization from the local education office to perform the study in the chosen middle schools during this visit. The selected middle school physics teachers were then trained.

Talking Physics is a set of educational materials prepared for training by physics education researchers at Addis Ababa University and Durham University. There are three types of booklets available from *Talking Physics*. The first type, *Talking Physics: In-Service Training*, introduces dialogic argumentation as a method of shifting physics education away from the traditional teacher-led didactic teaching, which makes students passive and limits their reasoning, and toward more student-centered dialogical teaching. In TPSS, students involved in dialogical argumentation were encouraged to critically analyze ideas using theoretical and empirical evidence. During the training, physics teachers examined the nature of physics education, student learning, and the physics curriculum, with a focus on upper primary physics. The training also included scientific reasoning, argumentation theory, and techniques to encourage dialogic argumentation in the classroom.

Talking Physics: Student Activities, the second booklet, covers 52 teaching activities, all of which were used in the training. All of these activities were carried out by physics teachers utilizing the pedagogy that we had urged them to employ in their physics classes. The final booklet, *Talking Physics: Teacher Guide,* advises physics teachers on how to modify their classroom instruction. It guides teachers through the process of incorporating dialogical argumentation into their classroom instruction. It also provides suggestions for time and group structure, as well as techniques for incorporating dialogical argumentation into their instruction. In addition, for each exercise, a correct response is provided.

During the training, we addressed the underlying ideas of all 52 activities, how to incorporate these activities into Grade 7 and 8 physics lessons, and the value of student talk in physics instruction. Four teacher education colleges coordinated and administered the three-day training. *Talking Physics: Student Activities* booklets were distributed to the intervention group at the end of the training.

Teachers used activities to measure students' scientific knowledge after introducing a topic or before offering a topic to bring forth students' scientific understanding. Teachers had also described the task to students, telling them what they needed to complete and how much time they had. In addition, teachers went around to different groups and offered questions to spark debate. Teachers encouraged students to discuss their views rather than simply providing the correct answer. Finally, in whole-class discussions, teachers summarized group conversations. Whole-class discussions were essential to first demonstrate why incorrect responses are incorrect and then to provide the correct responses.

9. Data collected

At two time points in the TPSS intervention year, data were collected using pre-tests and post-tests of physics reasoning test and video records of small group discussion tasks. The physics reasoning test instruments were developed for assessing students' physics knowledge and skills in scientific reasoning and argumentation, respectively. The quantitative data used included the pre-test and the post-test of the physics reasoning test to assess students' epistemic knowledge. To measure the quality of students' dialogical argumentation, we took 28 video recordings of middle school students' group tasks from 14 primary school classrooms. The group discussion used tasks that required students to use scientific reasoning and argumentation. The group was randomly selected and contained three to six middle school students. Overall, 14 video records were gathered for each of the activities in the treatment group and 14 records were gathered for each of the activities in the control group. Students were also encouraged to use any language that gives them a more comprehensive way of expressing their knowledge. All these video recordings were used in this study.

10. Data analysis

10.1 Inferential statistics used

We checked the assumption for normality, and we found that the students' pre-test and post-test scores of EKIs for both control and experimental groups were not normally distributed. Hence, we used the Mann–Whitney U test to compare the scores of students' pre-test and post-test scores of EKIs between control and experimental groups (between-subject comparison). We also computed the gains and normalized gains of both groups (within-subject comparison). Excel and SPSS (V.23) were used to analyze the data and visualize the findings using graphs. In addition, this study also explored the quality of dialogical argumentation during students' argumentative small group discussions using features of ASAC for both treatment and control groups.

10.2 Assessment of scientific argumentation in the classroom (ASAC) observation protocol

Many argumentation frameworks have been developed to analyze episodes of argumentation. ASAC was developed by Sampson, Enderle [41]. Sampson et al.'s main reason for the development of ASAC was to provide a tool "that can be used to assess the nature and quality of argumentation that occurs between students inside the science classroom" (pp. 236). ASAC situated argumentation as a process and targeted conceptual and cognitive, epistemic, and social aspects of argumentation. The conceptual and cognitive aspects of argumentation attempt to assess how students negotiate, make meaning, and develop scientific understanding. Seven ASAC items evaluate the level of in-depth scientific discussion, the use of alternative claims, the notification of discrepancy and the subsequent adjustment of claims or explanations, the level of challenge and negotiation to offered ideas, and the line of reasoning. ASAC's epistemic components of argumentation assess how students use evidence, theories, and models, how they evaluate evidence, and how they employ scientific languages in their discussions. The epistemic element was comprised of seven items of ASAC. These items assessed features of argumentation, such as how the content was presented, level of evidence usage, the relevance of evidence, quality of inference from data, how the argumentation is framed, and level of connectedness between inferences and observation. The social aspects of argumentation investigate the interaction. Five items examine the social nature of argumentation. This includes students' reflection of what they know and how they know, respect for ideas, willingness to contribute, and balanced engagement of group members.

The ASAC observation protocol consists of 19 items on a Likert scale that ranges from 0 (Not at all) to 3 (Often), only Item 6 and Item 8 were reversely rated, that is, from 3 (Not at all) to 0 (Often). Sampson and colleagues present an elaborated discussion about the theoretical underpinnings for ASAC and the methodological approaches they followed to confirm the construct validity, criterion validity, and reliability of the ASAC tool [41]. Comparison between ASAC and Toulmin's Arguments Pattern (TAP) framework considered ASAC as reliable in assessing the quality of students' argumentation, hence this study used ASAC to examine the quality of argumentation.

11. Results

This study analyzed the EKIs test scores and the ASAC scores to answer the extent and the ways dialogical argumentation affect middle school students' epistemic knowledge. The level of epistemic knowledge was determined using the students' scores in 10 EKIs. The overall physics reasoning was pilot tested and found to be valid and reliable (personal communication). The mean EKIs score then was compared progressively between the treatment and control group using SPSS.

11.1 Students' epistemic knowledge of science

Grade 8 students' epistemic knowledge of science was evaluated using 10 EKIs that were identified to measure epistemic knowledge. An independent sample t-test was used to determine if there were statistically significant differences in the students' epistemic knowledge levels over the course of the year. A high score in these 10 EKIs was regarded as a sign of superior epistemic knowledge. The pre-test and post-test mean scores for the treatment and control groups were used to determine the Grade 8 students' level of epistemic knowledge.

Figure 1 shows the results of this test. The treatment groups' level of epistemic understanding of science grew from 3.84 (±1.89) pre-test to 5.10 (±2.48) post-test, as measured by the mean test score (SD). The control groups' level of epistemic knowledge, on the other hand, was somewhat lower in the post-test (4.06 ± 2.35) than in the pre-test (4.14 ± 2.28). **Figure 2** shows that the mean score of students in the intervention group showed a slight increase in the post-test compared to their pre-test score. Not much difference was observed between the post-test score compared to their pre-test score for the control group. Additional statistical analyses were conducted to determine whether the mean EKIs score progressed significantly in the treatment group compared to the control group.

Students' scores were checked for statistical assumptions, such as normality, homogeneity, and outliers. Students' scores in the EKIs were first checked for outliers.

A boxplot assessment using SPSS identified seven outliers in the data for each of the control groups' pre-test and post-test scores and four outliers in the data for treatment groups' pre-test scores. Rather than removing the outliers, we substituted the values of these outliers with the next largest value (9.1 for control groups and 8.1 for treatment groups) plus small increments to maintain the ranking in the data. The distribution of students' scores on 10 EKIs appears to be not significantly different from a normal distribution, as evidenced by the histogram and normal Q-Q plot. However, Shapiro–Wilk's test ($p < 0.05$) revealed that the distribution of pre-test and post-test scores of both the experimental and control groups was not normal. Having obtained statistical evidence for non-normality, the nonparametric independent samples t-test was conducted to determine whether there was a significant difference between treatment and control groups in the pre-tests and post-tests of EKIs.

A Mann–Whitney U test was used to see if there was a statistically significant difference in the level of epistemic knowledge between the intervention and control groups based on their pre-test mean scores. The distribution (mean rank) of pre-test scores was not statistically significant between intervention and control groups, as indicated by the Mann–Whitney U test ($U = 27{,}126.5$, $z = 1.04$, $p = 0.299$, $r = 0.05$). This revealed that at the start of the intervention, Grade 8 students' epistemic knowledge of science was not substantially different among groups. However, for the post-test results, the Mann–Whitney U test ($U = 35{,}454$, $z = 4.5$, p 0.001, $r = 0.19$) revealed a statistically significant difference in EKI mean scores between the intervention and control groups. These small-to-medium effect sizes suggested a significant difference in post-test epistemic knowledge between the intervention and control groups. The implication is that engaging in dialogical argumentation has significantly increased the level of epistemic knowledge among Grade 8 students. The findings show that shifting physics instruction from didactic to dialogic may be beneficial.

Furthermore, Wilcoxon Signed Rank Tests were used to see if there was a statistically significant difference in the pre-test and post-test mean scores of both the intervention and control groups' epistemic knowledge. The level of epistemic knowledge in the control group was not significantly different on post-test scores than on pre-test

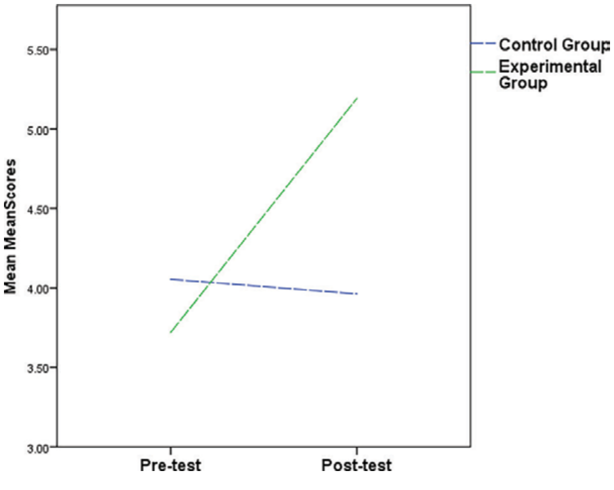


Figure 2.
Comparison of treatment and control group on a pre-test and a post-test of mean scores of the 10 epistemic knowledge items (n = 239 treatment group and n = 240 control group).

scores, T = 10,080, z = −0.208, p = 0.835, r = −0.01. However, a significant difference between pre-test and post-test scores was seen in the intervention group—level of epistemic knowledge was significantly higher on the post-test scores than on the pre-test scores, T = 17,220, z = 6.18, p = 0.000, r = 0.40. The medium to large effect sizes suggested a significant difference in the intervention group's level of epistemic knowledge between pre-test and post-test scores.

Table 1 presented the percentage of the average scores for EKIs before and after implementation of dialogical argumentation, the gain in epistemic knowledge based on EKIs scores, and the normalized gain. In physics education research, computing the normalized gain in studies that use pre- and post-tests are considered as a standard to evaluate curricula and instructional intervention. Normalized gain is the ratio of each student's improvement divided by capacity for improvement (Hake, 1998). Computation of normalized gains in this study reveals the improvement of students' epistemic knowledge of science between their pre-test scores and post-test scores. For the treatment group, where dialogical argumentation was used, the gain in epistemic knowledge, as evaluated by EKIs, was shown to be significant (8.40 percent gain) when compared to the control group (0.53 percent gain). Despite the fact that the students' EKIs scores were found to be very low (usually below the required minimum of 50%), we noticed a slight improvement in students' epistemic knowledge (a low gain of G = 0.11) for the treatment group compared to a gain (G = −0.01) for the control group.

11.2 Quality of students' dialogical argumentation

To assess the quality of students' dialogical argumentation, we recorded 28 small group discussions; 14 videos from before and after the intervention that engage them in dialogical argumentation for a year. The video records were collected from seven treatment and two control groups. Sampson, Enderle [41] developed an assessment tool, namely ASAC, which enhances the assessment of dialogical argumentation. The ASAC observation protocol considered conceptual, epistemic, and social aspects of argumentation as the fundamental construct during its development. Sampson, Enderle [41] believed that this tool would allow researchers to assess crucial aspects of argumentation, such as its nature and quality, as well as comprehensively analyze students' arguments.

The ASAC tool consists of 19 items, with seven examining conceptual aspects of argumentation, seven addressing epistemic aspects of argumentation, and the remaining five focusing on the social side of argumentation. Only Item 6 and Item 8 were reversely rated on a scale of 0 (Not at all) to 3 (Often). In this paper, the ASAC observation protocol was used to evaluate and compare the progress of Grade 8

Method of Teaching	Pre-test	Post-test	Gains	G	N
Dialogical Argumentation Classroom*	25.60	34.00	8.40	0.11	239
Traditional Instruction Classroom*	27.60	27.07	(0.53)	(0.01)	240

Both pre-test and post-test scores given here are found by changing the student's mean scores (given in* **Table 1*) into percentages.*

Table 1.
Average normalized gain of students' epistemic knowledge between argumentative and non-argumentative classrooms.

students across all elements of dialogical argumentation as a consequence of pedagogy intervention for a year. To establish the reliability of the ASAC rating of the episodes of argumentation, the first author and the second-rater, a physics teacher educator and PhD candidate in physics education at Addis Ababa University, assessed four randomly chosen episodes of student argumentative group discussion. The ASAC scorings have acceptable inter-rater reliability, as indicated by Cohen's kappa ($\kappa = 0.775$).

The ASAC observation protocol tool was used to rate 28 recorded student argumentative tasks. **Table 2** shows the overall ASAC mean scores for each school, as well as the mean scores for conceptual and cognitive aspects, epistemic aspects, and social aspects of argumentation.

Figure 3 (and **Table 2**) depicted the trends of change in the mean scores of different aspects of argumentation and the overall mean scores of ASAC between pre-test and post-test scores of both treatment and control groups. A critical look at the mean scores in **Table 2** and the time effect in **Figure 3** indicates that the overall quality of dialogical argumentation had shown progress in the treatment group compared to the control group. A Mann–Whitney test was used to determine if there were any significant differences in the quality of dialogical argumentation and various aspects of argumentation generated by the treatment and control groups as a result of their participation in argumentative and non-argumentative instructions, respectively. Here, all the effects are reported at $p < 0.05$. For the control group, the Mann–Whitney test revealed that there was no significant difference in the mean scores of Conceptual & Cognitive Aspects ($U = 17$, $r = -0.285$, $p = 0.285$), in the mean scores of Epistemic Aspects ($U = 11$, $r = -0.496$, $p = 0.064$), in the mean scores of Social Aspects ($U = 21.5$, $r = -0.853$, $p = 0.693$), and in the total mean ASAC scores ($U = 15$, $r = -0.332$, $p = 0.214$). For the treatment group, there was no significant difference in the mean scores of Conceptual & Cognitive Aspects ($U = 14.5$, $r = -0.35$, $p = 0.19$) and in the mean scores of Social Aspects ($U = 16$, $r = -0.299$, $p = 0.263$). Nevertheless, a statistically significant difference existed in the mean scores of Epistemic Aspects ($U = 5.5$, $r = -0.662$, $p = 0.013$) and in the total mean ASAC scores ($U = 8$, $r = -0.564$, $p = 0.035$).

11.3 Correlation between epistemic knowledge and dialogical argumentation

Pearson correlation was used to compare students' epistemic knowledge, as evaluated by mean scores on epistemic knowledge items, and the quality of their dialogical argumentation, as judged by the ASAC observation protocol. The correlation between the mean scores of the epistemic knowledge test and the mean scores of ASAC was done for 14 school cases. The Pearson's product–moment correlation analysis indicated a strong and positive correlation between mean scores of epistemic knowledge items and mean ASAC scores, $r(14) = 0.558^*$, $p = 0.038$ (*The correlation is significant at the 0.05 level (two-tailed)). Dialogical argumentation accounts for 31.14% of the variations in middle school students' epistemic knowledge of science.

Figure 4 provides a scatter plot between mean scores of epistemic knowledge based on mean scores of EKIs and mean scores of ASAC tool and its components of the aspect of argumentation. **Figure 4** also depicted the best fitting linear line and the proportion of variance described by the line ($R2 = 0.312$). The scatter plot in **Figure 4** describes the relationship between the development of epistemic knowledge and engagement in dialogical argumentation. This indicates that those who score high on ASAC also scored high in epistemic knowledge items, and vice versa.

Items	Item Descriptions and Dimensions of ASAC	Ctrl-Pre Mean(SD)	Ctrl-Post Mean(SD)	Intvn-Pre Mean(SD)	Intvn-Post Mean(SD)
Cca1	The conversation focused on the generation or validation of claims or explanations	0.43 (0.53)	0.71 (0.49)	0.71 (0.49)	1.43 (0.98)
Cca2	The participants sought out and discussed alternative claims or explanations	0.29 (0.49)	1(0.82)	1 (0.82)	1.14 (0.69)
Cca3	The participants modified their claim or explanation when they noticed an inconsistency or discovered anomalous information	- (–)	0.71 (0.49)	0.71 (0.49)	0.71 (0.49)
Cca4	The participants were skeptical of ideas and information	0.14(0.38)	0.43 (0.53)	0.43 (0.53)	0.57 (0.53)
Cca5	The participants provided reasons when supporting or challenging an idea	0.57(0.79)	1.14 (0.69)	1.14 (0.69)	1.57 (0.98)
Cca6	The participants based their decisions or ideas on inappropriate reasoning strategies	0.86(1.07)	2 (0.58)	2 (0.58)	1.86 (0.69)
Cca7	The participants attempted to evaluate the merits of each alternative explanation or claim in a systematic manner	0.29(0.49)	0.86 (0.69)	0.86 (0.69)	1.43 (0.98)
	Conceptual and Cognitive Aspects	**3.57**	**4.29**	**6.86**	**8.71**
Ea1	The participants relied on the “tools of rhetoric” to support or challenge ideas	2.29(0.95)	1.29 (0.49)	1.29 (0.49)	1.29 (0.49)
Ea2	The participants used evidence to support and challenge ideas or to make sense of the phenomenon under investigation	- (–)	0.71 (0.76)	0.71 (0.76)	1.43 (0.53)
Ea3	The participants examined the relevance, coherence, and sufficiency of the evidence	0.14 (0.38)	0.43 (0.53)	0.43 (0.53)	1.14 (0.69)
Ea4	The participants evaluated how the available data was interpreted or the method used to gather the data	0.29 (0.49)	0.43 (0.79)	0.43 (0.79)	0.86 (0.9)
Ea5	The participants used scientific theories, laws, or models to support and challenge ideas or to help make sense of the phenomenon under investigation	- (–)	0.29 (0.49)	0.29 (0.49)	0.86 (0.38)

Items	Item Descriptions and Dimensions of ASAC	Ctrl-Pre Mean(SD)	Ctrl-Post Mean(SD)	Intvn-Pre Mean(SD)	Intvn-Post Mean(SD)
Ea6	The participants made distinctions and connections between inferences and observations explicit to others.	0.14 (0.38)	0.57 (0.53)	0.57 (0.53)	1.14 (0.9)
Ea7	The participants used the language of science to communicate ideas.	0.14 (0.38)	0.57 (0.79)	0.57 (0.79)	1.57 (0.79)
	Epistemic Aspects	**3.00**	**4.29**	**4.29**	**8.29**
Sa1	The participants were reflective about what they know and how they know.	1 (–)	1.14 (0.38)	1.14 (0.38)	1.43 (0.53)
Sa2	The participants respected what each other had to say.	1.43 (0.53)	1.43 (0.53)	1.43 (0.53)	1.57 (0.53)
Sa3	The participants discussed an idea when it was introduced into the conversation.	0.43 (0.53)	1.14 (0.69)	1.14 (0.69)	1.86 (0.38)
Sa4	The participants encouraged or invited others to share or critique ideas.	0.43 (0.53)	1.43 (0.53)	1.43 (0.53)	1.57 (0.53)
Sa5	The participants restated or summarized comments and asked each other to clarify or elaborate on their comments.	0.57 (0.79)	1 (0.58)	1 (0.58)	1.14 (0.69)
	Social Aspects	**3.86**	**3.57**	**6.14**	**7.57**
	ASAC Observation Protocol Total Score	**10.43**	**12.14**	**17.29**	**24.57**

Table 2.
Mean ASAC scores of both treatment and control groups before and after a yearlong intervention in dialogical argumentation.

12. Discussion

This quasi-experimental study aimed at finding the impact of a yearlong engagement in dialogical argumentation on Grade 8 students' epistemic knowledge of science. The quantitative results presented above showed an overall improvement in epistemic knowledge. In addition, analysis of video recordings of 28 episodes of argumentation revealed improvements in the experimental groups in the quality of dialogic argumentations. The improvements in epistemic understanding of science and the quality of dialogic argumentation were found to be linearly correlated.

Two research questions guided this study. The main purpose of the first research question was to explore Grade 8 students' prior epistemic knowledge and then to investigate whether learning physics through dialogical argumentation would change their epistemic understandings. The results indicated that both experimental and control groups show low levels of epistemic knowledge. After the intervention, the experimental groups showed a small but significant improvement in their level

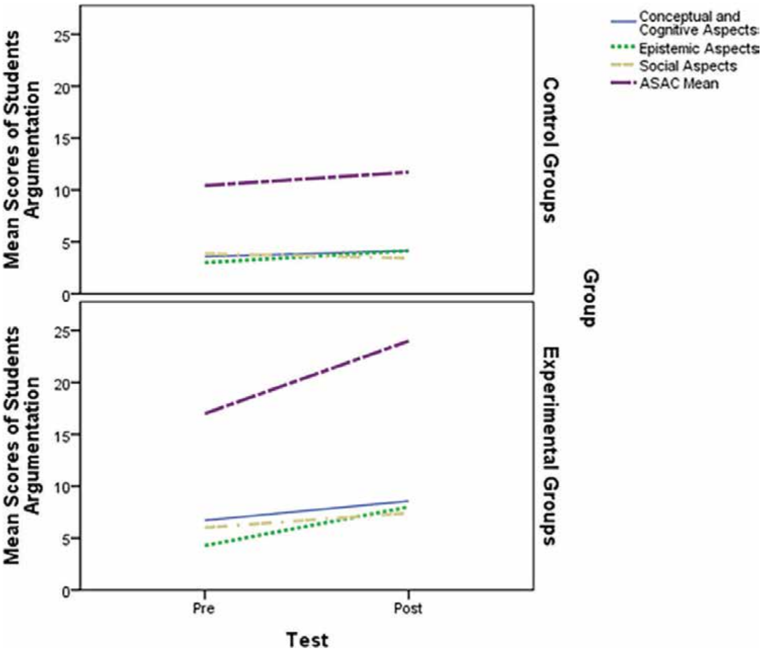


Figure 3.
Differences in the level of conceptual and cognitive, epistemic, social aspects of argumentation, and the overall ASAC scores between treatment and control groups across time (pre-test and post-test).

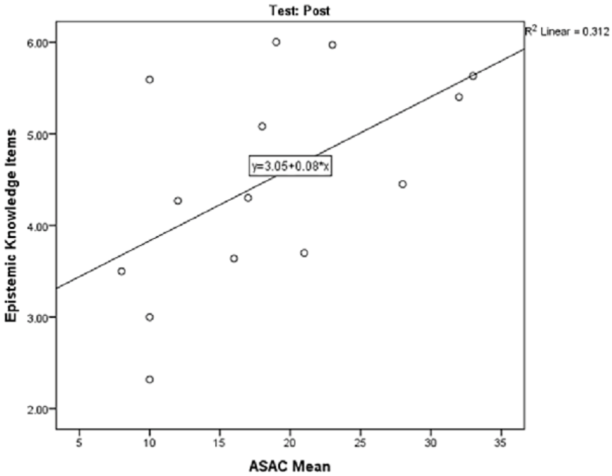


Figure 4.
A scatter plot of mean scores of epistemic knowledge items and mean overall scores of ASAC observation protocol.

of epistemic understanding compared to the control group. These findings suggested that dialogic argumentation did facilitate the growth of students' epistemic understanding of science. Similar positive effects of argumentation on students' epistemic cognition have been reported in the literature [8, 42, 43]. That is, this result is consistent with previous findings that indicated that argumentation could provide a powerful context for improving students' epistemic understanding of science [44, 45]. The main purpose of the second question was to investigate whether continuous engagement in dialogic argumentation would improve the quality of student-generated dialogic argumentations and whether the level of epistemic understanding would correlate with the quality of dialogic argumentation. In this study, no explicit instruction was given on how to improve the quality of student argumentation, though we had trained elementary school teachers about dialogical argumentation.

The quality of students' dialogic argumentations remained very low and unchanged for the conceptual and cognitive aspects of dialogic argumentation and social aspects of dialogic argumentation. However, though the quality of the epistemic aspect of argumentation was still low, statistically significant scores were observed in favor of the experimental group compared with the control group. This was mainly due to students' poor performance with regard to Item 3 and Item 4 of the conceptual and cognitive criteria and with regard to Item 10, Item 11, and Item 12 of the epistemic criteria in the ASAC protocol. They had the tendency to use commonsense reasoning to support or challenge claims rather than draw on the available evidence and principles, laws, theories, and formulae (Item 10). The second problem was the naïve use of evidence, that is, no or minimal attempts to examine its relevance, coherence, and sufficiency to the wave phenomenon being discussed (Item 10), and the third was making vague inferences (Item 13). Though the Mann–Whitney U test indicated that there was a statistically significant difference in the epistemic aspects of argumentation between experimental and control groups, the raw scores were still low. These results meant that Grade 8 students' understanding of the way of knowing science and their skills for arguing with supportive evidence and rationale were at a low level. Other researchers also noted that students are unfamiliar with the "norms of scientific argumentation" and misunderstand what counts as good evidence and reasoning in science (e.g., Walker and Sampson [10], Sampson and Clark [46], Simon, Erduran and Osborne [47]) reported similar problems.

Generally, the quality of dialogic argumentation improved. This was expected because dialogic argumentation is a cognitive process that needs the practice to understand the epistemological foundations of science and to develop the ability to reflect on theories and evaluate them using evidence [3, 48, 49]. Nonetheless, this result suggests that sustained engagement in dialogical argumentation has a positive impact on argument quality. This is consistent with Osborne and Erduran's [34] observation that under adequate instructional settings, students can considerably enhance their ability to argue.

A few studies have also explored the relations between epistemologies and particular scientific practices, such as argumentation [42, 43, 50]. These studies generally focus on the impact of epistemic knowledge on students' argumentation abilities. For example, Mason and Scirica [43] found that the quality of students' arguments, counterarguments, and rebuttals correlated with their level of epistemological sophistication. The current study also found a positive or direct association between students' epistemic knowledge and the quality of their argumentations.

There are some limitations to the current study. The first is the random selection of the groups who participated in a video-recorded small group discussion task. The selected group may not be representative of the classroom. Consequently, it may not depict the complete picture of the classroom. Being aware of this limitation, we had informed teachers to compose a group representative of the classroom. The teachers' verdicts were used to the trustworthiness of the selected group's representation. The other limitation could be the influence of instructional approaches used in other school subjects. In this yearlong study, only physics teachers used dialogical argumentation, whereas biology and chemistry teachers were free to use any instructional approach. Thinking that biology and chemistry are part of science, there may be situations that affect the Grade 8 students' development of argumentation skills and their epistemic knowledge.

To summarize, this study clearly reveals improvements in the level of student argumentation ability as a result of yearlong participation in dialogical argumentation. Similarly, the study demonstrates not only the existence of a positive correlation between students' level of epistemic knowledge and the quality of their dialogic arguments but also that the correlation was strong. Though the relationship between student argumentation and their epistemic knowledge remains controversial, the outcomes of this study give additional evidence for the presence of a positive association between epistemic knowledge and the quality of dialogic argumentation.

13. Conclusions

Extensive searches for similar findings in relevant science education and learning science literature revealed that the effect of students' engagement with dialogic argumentation on epistemic knowledge or vice versa appeared to be understudied so far, both in elementary and secondary school physics and in physics teacher training settings. Furthermore, to the best of the researchers' knowledge, there is no evidence that the quality of dialogic argumentation is linked to the development of epistemic knowledge. Using dialogic argumentation to improve Grade 8 students' epistemic knowledge and matching epistemic practices in science classes with the epistemic practices of the scientific community has been a necessary but difficult job [15].

This study included 14 upper primary schools from Ethiopia's three regional states (Addis Ababa, Amhara, and South Peoples). The research was conducted as part of the TPSS project, which ran from 2016 to 2018. Data were gathered by administering physics reasoning tests, which included epistemic knowledge items (EKIs), video recordings of students' small group discussions, video recordings of whole-class teaching, and audio recordings of teacher interviews before and after the intervention. Teachers in the intervention groups received a three-day training in dialogical argumentation. During the training, these teachers were exposed to the Talking Physics manual (created by the TPSS project), with follow-up talks on how to implement the manual's activities.

This study's results indicated that dialogical argumentation was helpful in increasing Grade 8 students' epistemic knowledge in physics as well as their competence to reason scientifically during arguments. The significant correlation between students' level of epistemic knowledge and progress in the quality of their dialogical argumentation would imply that dialogical argumentation indeed helps Grade 8 students improve their epistemic knowledge in physics. We statistically compare the mean scores of EKIs for the two groups and observe that there are no plausible reasons why the means of the two groups might be different, other than the possibility that the argumentative instructions do have a differential impact on students' epistemic knowledge of science compared to the non-argumentative instruction.

The findings from this study have implications for the teaching of science at middle schools. In most middle schools in Ethiopia, science lessons are often taught through lectures and demonstrations. Much emphasis is given to verifying scientific laws and theories, that is, the content and conceptual aspect of scientific knowledge. This suggests that little attention is given to engaging students in a discursive interaction to construct scientific knowledge. Therefore, physics teachers should provide an opportunity to their students to entertain competing views of a certain physical phenomenon when teaching physics and should encourage students' dialogical discourse. In this study, Grade 8 students demonstrated small but significant improvement in their epistemic knowledge after they had been exposed to dialogical argumentation compared to other Grade 8 students who were not exposed to dialogical argumentation. The low level of both groups of students' epistemic knowledge of science in the pre-test indicates that their physics lessons do not adequately address and not well integrated the ways of knowing science. Nevertheless, the small and significant improvement of students' epistemic knowledge in the experimental group because of dialogical argumentation lessons reveals that argumentation is a viable approach to equip students about the knowledge generation and construction mechanism in science. This implies that middle school physics teachers could scaffold their students' epistemic understanding of science by shifting their teaching toward a dialogical pedagogy. Engaging students in dialogical argumentative tasks will provide the necessary experiences to develop and master the epistemic entities of scientific reasoning. This is also consistent with global efforts to promote student-centered practices through methodological shifts in science instructions. It should also be noted that the positive results in the current study were obtained even though there is a considerable emphasis on teaching through transmission in Ethiopia. The results of this study corroborated prior findings that the transmission mode of teaching is ineffective in promoting science teaching and learning and should be replaced with student-centered instructional strategies, such as dialogic argumentation. Students can use dialogical argumentation to develop an epistemic understanding of the role of scientific evidence and how science operates. As a result, students will be better prepared to make informed decisions concerning various scientific knowledge claims and their applications.

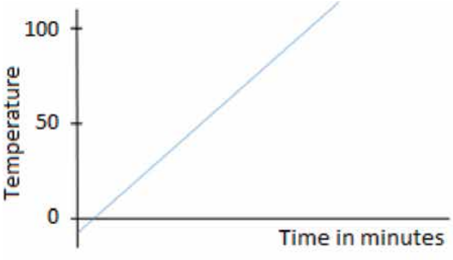
In a summary, this research revealed that students' participation in argumentative instructions significantly improved their epistemic knowledge and argumentation quality. Furthermore, the study found that a significant change in the quality of argumentation, as measured by ASAC scores, is associated with a significant change in students' ASAC epistemic scores. Other areas of ASAC, such as conceptual and cognitive aspects, and social aspects, showed no significant changes.

A.Epistemic knowledge items from physics reasoning test

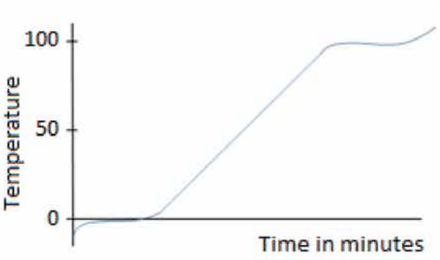
Item 3

Scientists claim the temperature does <u>not</u> change when ice is melting or when water is boiling.

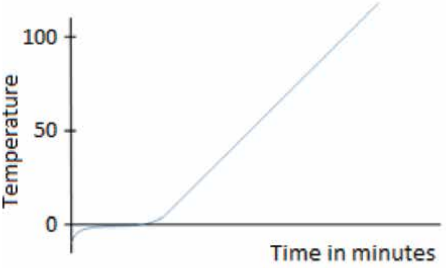
Which graph below supports the scientists' claim?



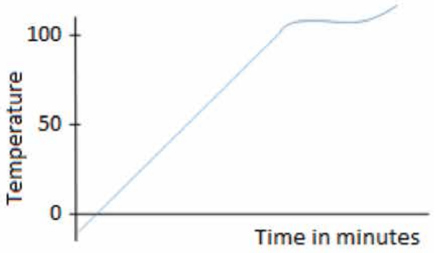
A



B



C



D

Item 4
Scientists say matter can turn from **solid** to **liquid**.
How do we know they are right?

A. Because ice can melt to water

B. Because ice can be crushed to powder

C. Because ice expands when it freezes

D. Because ice can be hard as rock

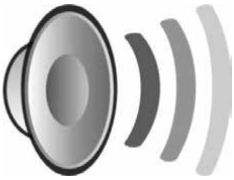
Item 5
Scientists say sound is caused by **vibrations**
How do we know they are right?

A. Because vibrations can make things break

B. Because you can feel vibrations in the strings of a guitar or a kirar

C. Because we can hear sound coming from a guitar or a kirar

D. Because loud sound can make the ears hurt



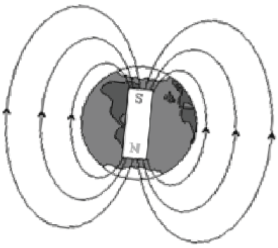
Item 6
Scientists say the Earth has a **magnetic field**.
How do we know they are right?

A. Because we can find magnetic stones in nature

B. Because we can see magnets attract each other

C. Because we can see that the Sun attracts the Earth

D. Because we can see the needle on a compass always pointing to the North



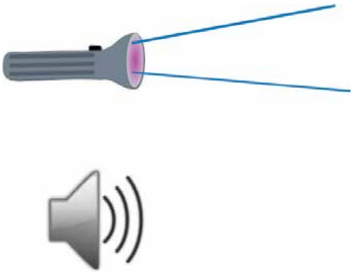
Item 7
Some students think sound travels **faster** than light.
How do we know they are <u>wrong</u>?

A. Because you can see a lightning before you can hear the thunder

B. Because light travels all the way from the Sun to the Earth

C. Because you can still hear sounds when it is dark

D. Because sound travels faster in water than in air



Item 8

Some students think the Moon, like the Sun, is producing its own light.
How do we know they are wrong?

A. Because the Moon is seldom seen at daytime

B. Because we feel no heat from the Moon

C. Because shadows caused by moonlight are very weak

D. Because the part of the Moon not lit by sunlight is dark

Item 9

Debre thinks boats float because they are made of light material.
Decide if the observations below support or oppose what Debre thinks (**tick one box for each observation**)

Observation	Supports Debre	Oppose Debre
When you drop a stone in water it sinks, but a piece of wood floats	☐	☐
Some boats are made of steel, but still float	☐	☐
Some stones are very light, but still sinks	☐	☐

Item 10

The table shows average temperature at different places in Ethiopia

Location	Average temperature (°C)	Altitude (meters above sea level)
Arba Minch	21.8	1278
Dilla	20.6	1572
Bahir Dar	19.6	1797
Hawassa	19.2	1710
Gondar	19.1	1966
Addis Ababa	15.9	2324
Debre Markos	15.9	2462
Dessie	15.2	2491

a. What is the average temperature in Gondar________________°C?

b. How does temperature change with altitude?

A. Higher altitude has higher temperature

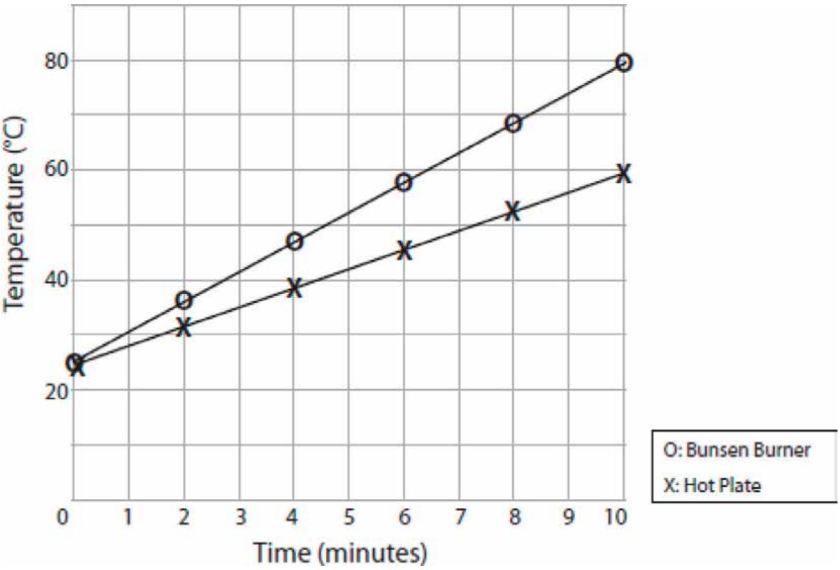
B. Higher altitude has lower temperature

C. There is no clear pattern between altitude and temperature

Item 11

Some students investigated two ways of heating water. One way was hot plate and the other was Bunsen burner.

They heated the same amount of water and measured temperature every 2 minutes. The measurements are presented in the graph:



a) Which heat source (hot plate or Bunsen burner) heated the water fastest?

____

b) What was the difference in temperature after 10 minutes?

____

Item 12

The table compares properties of three materials: **Wood, rock** and **iron**

Property	Material 1	Material 2	Material 3
Sinks in water?	Yes	No	Yes
Burns easily?	No	Yes	No
Attracted by a magnet?	Yes	No	No

Decide which materials are wood, rock and iron (write the number)
Wood is material number ____________
Rock is material number ____________
Iron is material number ____________

Author details

Getachew Tarekegn[1]*, Jonathan Osborne[2] and Mesfin Tadesse[3]

1 Department of Natural Sciences, Gondar College of Teachers Education, Gondar, Ethiopia

2 Graduate School of Education, Stanford University, Stanford, USA

3 Department of Physics, Addis Ababa University, Addis Ababa, Ethiopia

*Address all correspondence to: teobista@gmail.com

References

[1] Lehesvuori S. *Towards Dialogic Teaching in Science: Challenging Classroom Realities through Teacher Education (Unpublished doctoral dissertation), in Department of Teacher Education*. Jyväskylä: University of Jyväskylä; 2013

[2] OECD. Draft PISA 2015 Science Framework. Paris, France: OECD Publishing; 2013

[3] Duschl RA. Science education in three-part harmony: Balancing conceptual, epistemic, and social learning goals. Review of Research in Education. 2008;**32**(1):268-291

[4] Kind P, Osborne J. Styles of scientific reasoning: A cultural rationale for science education? Science Education. 2017;**101**(1):8-31

[5] Osborne J, Rafanelli S, Kind P. Toward a more coherent model for science education than the crosscutting concepts of the next generation science standards: The affordances of styles of reasoning. Journal of Research in Science Teaching. 2018;**55**(7):962-981

[6] Iwuanyanwu PN, Ogunniyi MB. Effects of dialogical argumentation instructional model on pre-service teachers' ability to solve conceptual mathematical problems in physics. African Journal of Research in Mathematics, Science and Technology Education. 2020;**24**(1):129-141

[7] Mhakure D, Otulaja F. Culturally-responsive pedagogy in science education: Narrowing the divide between Indigenous and scientific knowledge. In: Otulaja FS, Ogunniyi MB, editors. The World of Science Education: Handbook of Research in Science Education in Sub-Saharan Africa. Rotterdam, The Netherlands: Sense Publishers; 2017

[8] Murphy PK et al. Fostering high school students' conceptual understanding and argumentation performance in science through quality talk discussions. Science Education. 2018;**102**(6):1239-1264

[9] von Aufschnaiter C et al. Arguing to learn and learning to argue: Case studies of how students' argumentation relates to their scientific knowledge. Journal of Research in Science Teaching. 2008;**45**(1):101-131

[10] Walker J, Sampson V. Learning to argue and arguing to learn: Argument-driven inquiry as a way to help undergraduate chemistry students learn how to construct arguments and engage in argumentation during a laboratory course. Journal of Research in Science Teaching. 2013;**50**(5):561-596

[11] Stroupe D, Moon J, Michaels S. Introduction to special issue: Epistemic tools in science education. Science Education. 2019;**103**(4):948-951

[12] Woodward DH. Teaching science in ethiopia. American Scientist. 1969;**57**(4): 338A-344A

[13] Erduran S, Dagher Z. Reconceptualizing the Nature of Science for Science Education: Scientific Knowledge, Practices and Other Family Categories. Dordrecht, The Netherlands: Springer; 2014

[14] Manz E. Representing student argumentation as functionally emergent from scientific activity. Review of Educational Research. 2015;**85**(4):553-590

[15] Berland LK et al. Epistemologies in practice: Making scientific practices meaningful for students. Journal

of Research in Science Teaching. 2016;**53**(7):1082-1112

[16] Öberg G, Campbell A. Navigating the divide between scientific practice and science studies to support undergraduate teaching of epistemic knowledge. International Journal of Science Education. 2019;**41**(2):230-247

[17] Matthews M. Science Teaching: The Contribution of History and Philosophy of Science, 20th Anniversary Revised and Expanded Edition. 2nd ed. New York, NY: Routledge; 2015

[18] Duschl RA. Quality argumentation and epistemic criteria. In: Erduran S, Jimenez-Aleixandre MP, editors. Argumentation in Science Education: Perspectives from Classroom-Based Research. Dordrecht, The Netherlands: Springer; 2008:159-175

[19] Yang F-Y et al. High-school students' epistemic knowledge of science and its relation to learner factors in science learning. Research in Science Education. 2018;**48**(2):325-344

[20] Nielsen JA. Dialectical features of students' argumentation: A critical review of argumentation studies in science education. Research in Science Education. 2013;**43**(1):371-393

[21] Tadesse M. Dialogic Argumentation as Pedagogy: Impact on Preservice Physics Teachers' Conceptual Understanding of Mechanical Waves A Mixed Methods Study (Unpublished doctoral dissertation) in SME. Addis Ababa, Ethiopia: Addis Ababa University; 2015. p. 356

[22] Lehrer R, Schauble L. The development of scientific thinking. In: Liben L, Muller U, Lerner R, editors. Handbook of Child Psychology and Developmental Science. Hoboken, NJ: Wiley; 2015. pp. 671-714

[23] Iordanou K, Constantinou CP. Developing pre-service teachers' evidence-based argumentation skills on socio-scientific issues. Learning and Instruction. 2014;**34**: 42-57

[24] Cavagnetto A. Argument to foster scientific literacy: A review of argument interventions in K-12 contexts. Review of Educational Research. 2010;**80**(3):336-371

[25] Nussbaum EM. Collaborative discourse, argumentation, and learning: Preface and literature review. Contemporary Educational Psychology. 2008;**33**(3):345-359

[26] MoE & JICA. In: Education FMo, editor. Strategic Policy for National Science, Technology and Mathematics Education. Addis Ababa, Ethiopia: MoE & JICA; 2016

[27] NRC. A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. Washington, DC: National Academies Press; 2012

[28] Mijung K, Wolff-Michael R. Dialogical argumentation and reasoning in elementary science classrooms. Leiden, The Netherlands: Brill; 2019

[29] Mercer N. Developing argumentation: Lessons learned in the primary school, in argumentation and education: Theoretical foundations and practices. Nathalie M, Anne-Nelly PC, editors. Dordrecht: Springer; 2009: 177-194

[30] Alexander RJ. Towards Dialogic Teaching: Rethinking Classroom Talk. 4th ed. York: Dialogos; 2008

[31] Reznitskaya A, Gregory M. Student thought and classroom language: Examining the mechanisms of change in dialogic teaching. Educational Psychologist. 2013;**48**(2):114-133

[32] Kelly GJ, Licona P. *Epistemic practices and science education, in History, philosophy and science teaching: New perspectives*, Maththews MR, editor. Cham, Switzerland: Springer; 2018. p. 139-165

[33] Kelly G. *Inquiry, activity and epistemic practice, in Teaching Scientific Inquiry*. Duschl R, Grandy R, editors. Leiden, The Netherlands: Brill; 2008. pp. 99-117

[34] Osborne J, Erduran S, Simon S. Enhancing the quality of argumentation in school science. Journal of Research in Science Teaching. 2004;**41**(10):994-1020

[35] Iordanou K, Constantinou CP. Supporting use of evidence in argumentation through practice in argumentation and reflection in the context of SOCRATES learning environment. Science Education. 2015;**99**(2):282-311

[36] Kuhn D et al. Developing norms of argumentation: Metacognitive, epistemological, and social dimensions of developing argumentive competence. Cognition and Instruction. 2013;**31**(4):456-496

[37] Creswell J, Clark V. Designing and Conducting Mixed Methods Research. 3rd ed. Los Angeles: Sage; 2018

[38] Bryman A. Social Research Methods. 4th ed. Oxford: Oxford University Press; 2012

[39] Rapanta C, Felton M. Mixed methods research in inquiry-based instruction: An integrative review. International Journal of Research & Method in Education. 2019;**42**(3):288-304

[40] MoE. In: Education Mo, editor. Curriculum Framework for Ethiopian Education (KG - Grade 12). Addis Ababa, Ethiopia: Ministry of Education; 2009

[41] Sampson V, Enderle P, Walker J, *The Development and Validation of the Assessment of Scientific Argumentation in the Classroom (ASAC) Observation Protocol: A Tool for Evaluating How Students Participate in Scientific Argumentation, in Perspectives on Scientific Argumentation: Theory, Practice and Research*. Khine M, editor. Springer: Dordrecht; 2012:235-264

[42] Greene JA, Seung BY. Modeling and measuring epistemic cognition: A qualitative re-investigation. Contemporary Educational Psychology. 2014;**39**(1):12-28

[43] Mason L, Scirica F. Prediction of students' argumentation skills about controversial topics by epistemological understanding. Learning and Instruction. 2006;**16**(5):492-509

[44] Kuhn D et al. Can engaging in science practices promote deep understanding of them? Science Education. 2017;**101**(2):232-250

[45] Osborne J. Arguing to learn in science: The role of collaborative, critical discourse. Science. 2010;**328**(5977):463-466

[46] Sampson V, Clark D. Assessment of the ways students generate arguments in science education: Current perspectives and recommendations for future directions. Science Education. 2008;**92**:447-472

[47] Simon S, Erduran S, Osborne J. Learning to teach argumentation: Research and development in the science classroom. International Journal of Science Education. 2006;**28**(2-3):235-260

[48] Kuhn D. Teaching and learning science as argument. Science Education. 2010;**94**(5):810-824

[49] Sandoval W. Understanding students' practical epistemologies and their influence on learning through inquiry. Science Education. 2005;**89**(4):634-656

[50] Ryu S, Sandoval WA. Improvements to elementary children's epistemic understanding from sustained argumentation. Science Education. 2012;**96**(3):488-526

Chapter 3

Views of South Sudanese Secondary School Teachers about the Use of Humour in the Mathematics Classroom

William Deng Tap, Helicopter Mark Bulbul
and Biar Simon Ajang

Abstract

This chapter reports the views of South Sudanese secondary school teachers about the use of humour in the mathematics classroom as a teaching and learning tool. The use of humour as a pedagogical toolkit in a mathematics classroom is something that has not yet been seriously or widely considered and how the teachers, especially South Sudanese teachers, would react to the use of humour in the classroom was not yet known. An opinion survey containing six (6) close-ended questionnaire items or statements related to the use of humour in the classroom was distributed to ten (10) secondary schools located within and around Juba city. About sixty-five (65) South Sudanese secondary school teachers responded to the survey. Posed was a research question intended to explore the general views, attitudes, or opinions of South Sudanese secondary school teachers: What do South Sudanese secondary school teachers think about the use of Humour-Supported Instructional Approach (H-SIA), a proposed-alternative method of teaching and learning mathematics at secondary school level? Findings of this opinions survey indicate that South Sudanese secondary school teachers' overall average views are positive toward the use of humour in the classroom setting. The average majority of the surveyed secondary school teachers appeared keen and seemed eager to welcome experimentation with new ways of teaching and learning in the classroom. Hence, it is recommended that classroom teachers be always encouraged and allowed a certain degree of freedom to explore and try out new ways of teaching and learning. It is suggested, however, that teachers be first provided with necessary proper training about how to use humour appropriately, effectively, and creatively in the classroom environments.

Keywords: teachers' views, classroom humour, mathematical humour, pedagogical toolkit, teaching and learning mathematics

1. Introduction

This chapter reports an aspect of mathematics teachers' views about the use of humour in mathematics classroom that emerged during the pilot phase of a larger study on a newly proposed method of instruction called Humour-Supported Instructional Approach (H-SIA) [1]. In such a study [1] where mathematics itself was described or perceived as an infinite task, process, or procedures performed by a finite person, performer, or operator, the purpose of the method (H-SIA) was to attract and enhance students' interest in mathematics using the concept of humour as a teaching and learning tool. Viewed from this perspective or perception of mathematics as an infinite process with an infinite number of objects to operate with (e.g., it may take nearly an infinite amount of mathematical operations or procedures just to figure out what is happening in a tinny place such as an open interval on a line or a surface of a tinny disc or ball), it is not difficult to see why there always been legitimate issues, concerns or problems associated with the teaching and learning of mathematics as subject matter. Hence, most students always privately complain about the difficulties involved in learning, teaching and even mastering such an infinite task, procedures, or infinite operations; although some students sometimes openly complain but often in the guise of boredom (lack of interest) in various mathematics classroom settings, perhaps fearful of the negative stereotypes associated with the illusive labels of the so-called IQ concept or myth. Dealing with an infinite task is always a delicate business or issue and it is therefore not surprising why we always hear various students' stories about mathematics classrooms that are so boring, about mathematical experiences that are so humiliating, or about mathematics tasks that are so pointless [2]. Hence, it pays to lighten up the subject with humour or in other ways that are appreciated by the students [3, 4] as humour itself is known to play an important role in people's overall well-being in their daily-life activities [5].

Humour as a teaching technique or strategy has been widely defined as anything perceived or recognised by students to be funny, comical, or amusing; or the quality that makes something funny, laughable or amusing [6, 7]; and is more recently generally defined as an skillset, away to communicate, an educational strategy, a personal perspective or a positive emotional and behavioral response [8]. In this study, however, humour was specifically defined relative to mathematics content area as "mathematical humour" [9], a mathematics content-related humour often derived from mathematical concepts being discussed, combined with general humour ideas, particularly the incongruity theory of humour characterised by elements of surprises and unexpected twists or turns [5]. This opinions survey about teachers' views or attitudes was motivated by a South Sudanese volunteer mathematics teacher who initially agreed to take part as a co-teacher-researcher during the experimentation and implementation phase of H-SIA [10], but who later opted out of research participation. This left the researchers to wonder whether the teacher's apparent quitting has anything to do with South Sudanese teachers' overall general views, attitudes or opinions toward the use of humour in the classroom as a teaching and learning tool. While the mostly-welcoming views of students-learners toward the use of humour in the classroom settings are well documented [11, 12], the general views, opinions, or attitudes of the classroom teachers themselves are not yet widely explored [13–15]. The researchers then posed a guiding question intended to explore the general views, attitudes, or opinions of South Sudanese secondary school classroom teachers: What do South Sudanese secondary school teachers think about the use of H-SIA? The H-SIA was a proposed-alternative method of teaching and learning mathematics at the secondary school level for displaced South Sudanese students living in re-settled communities.

The purpose of the proposed method (H-SIA) was to generate and maintain interest in mathematics for these types of students. These students were among the internally displaced people living in temporary resettlement camps or shelters called Protection of Civilians sites (POCs) under the protection of the United Nation Mission in South Sudan (UNMISS). The students were displaced by the South Sudanese civil war and were taught in five secondary schools, namely Hope, Mat (Union) 1 & 2, Future, and Equity Senior Secondary Schools. These secondary school students were taught by a group of volunteer teachers who double as primary school teachers in adjacent UNICEF sponsored feeder primary schools. These students were chosen for this study because of their exposure to traumatic situations such as forced displacement, committed atrocities, poverty, and living in an insecure environment characterised by threats and intimidations. These students were perceived to be more concerned with issues of daily life activities for survival rather than the learning of mathematics in the classroom setting.

2. Background for the study

While the role of humour in the classroom appears to have been generally ignored, which implies that its use in the classroom is often rare if not almost none existence [4, 15–26], humour in its explicit-verbal forms such as storytelling or narrative has been used almost from the beginning of human existence as social beings. Storytelling technique, funny-humorous stories or narratives [10, 23, 27] as an expressive-verbal form of humour has been used since time immemorial as a powerful tool for grabbing the audience's attention and for passing on knowledge from generation to generation, culture to culture, and individual to individual. Even today, with the invention of powerful technological tools such as PowerPoint presentations, overhead projectors, chalkboards, or even the highly hyped smart boards, storytelling or narrative as a verbal form of humour remains an effective and persuasive channel for knowledge transmission [14].

Storytelling, funny-humorous stories, or exciting narratives [5, 28–30] as an expressive-verbal form of humour is one of the many old known classroom techniques and strategies, but one that seemed to be used only by very few passionate, creative and effective classroom teachers to motivate or inspire their students in the learning process [31–34]. Using humour as a teaching and learning technique or strategy [16] is necessary even more so in an environment such as the mathematics classroom where things tend to be run in a machine-like-robotic fashion with little or no regard to human feelings or emotions. Therefore, classroom humorous materials such as instructional humour [35], storytelling, stories or narratives [36–39] along with many other appropriate various types of humour can be creative way in humanising mathematics and the mathematics classroom [40–42]. Instructional humour, that is, humour related to content material such as mathematical humour [43–51] can be useful and effective in motivating and inspiring students to learn. While mathematical humour can convey to students the teaching qualities such as the teacher's authentic passion for the subject, enthusiasm, and curiosity for teaching and learning, it can also help in fostering not only the much desired cognitive factors but also the often neglected affective learning domains of the subject that are essential for student's social development.

Somehow explicit yet appropriate-contextualised examples of mathematical humour (right off the bat) for secondary school context would be a thief, a shopkeeper, and the total amount stolen: In this case, a thief first stole a 100 dollars from your shop and then came back moments later and bought stuffs worth 75 dollars and you gave him back a 25 dollars change, what is then the total amount of money stolen from you? A chicken and half that lays one and a half eggs a day: If a chicken and half lays one and half eggs a day, what is then the total number of chickens after certain number of days, months, years, and even at eternity? Or the famous hypothetical classic example of the identical pair of twins paradox in physics (special relativity) used to illustrate that the attractive concept of simultaneity is nothing more than as a matter of converged-diverse personal opinions, where one of the twins (Albert) is supposedly born in New York while the other twin (Alvin) in Tokyo: This has to be admittedly one of those hilarious-classic grand mothers' borderline humour but appropriately contextualised jokes where the twins' mother is supposedly so huge and gigantic such that she is more than capable of delivering these two-cute-identical babies at the same time, but at different places further apart. Here, a contextualised-mathematics problem for the secondary school level would be figuring out the total distance or separation, given the average speeds for the first and second halves of the journey between the two cities; and after that, the most curious students would then be challenged to derive the familiar half way formula from the total separation or distance between the two cities (New York and Tokyo). For more detailed discussion of various examples of mathematical humour relative to the content materials being discussed or the typical appropriate-classroom humour, readers can refer to the following related articles: [1, 4, 6, 12, 14, 18, 25, 50, 52]. In fact, there are many-various opportunities out there for wonderful stories or funny-humorous stories related to mathematics that are appropriate for secondary school contexts and beyond such as, for example:

> *Decartes' proof of existence of God, Pascal's famous wager, Plato's world of forms, and Newton's attempt to verify biblical chronology, Liebniz's detailed theodicy, current attempts to describe a divine domain in terms of meta-system, and mystical speculation on the infinite. ([45], pp. 62–63).*

Instructional humour or classroom humour [24] such as mathematical humour [1, 5] is humour derived from the mathematics concepts being discussed combined with general humour ideas, particularly the incongruity-resolution theory of humour often characterised by elements of surprises, exaggerations, unexpected twists or turns [22, 30–34, 38, 51, 53]. Infused and laced into H-SIA lessons as a form of related humour are pseudo-mathematical proofs or funny-humorous demonstrations–such as, for example, the latest-recent mathematical attempt, claim, or partial proof for the existence of some sort of "the real part of God," although not yet its complete, complex or whole part [35], a reminiscent of such similar attempts in the past into mathematical theology adventures by the legends before us–attempts which are more often than not riddled with hidden commonly made mistakes and/or ambiguous-unexamined tacit assumptions (e.g., division or multiplication by a disguised zero) which fall into these categories of incongruous elements of surprises, irony, exaggerations, unexpected twists or turns. However, according to Weber [22], false proofs or demonstrations are "funny only if there is some interpretation scheme by which they could plausibly make sense" (p. 58). In addition, all the mathematical humour used in the classroom setting must be either below and/or contextualised at the level of mathematics concepts being taught so that such mathematics jokes are well understood by the students-learners [1].

The primary aim of using humour as an instructional approach in mathematics classroom is to lighten or fire-up students' learning experience, inspire or motivate students to develop liking (interest) for the subject matter, and perhaps foster students' social skills students' social development is an aspect that is often ignored in the mathematics classroom, but one that is equally as important as student's cognitive development often promoted in such environment.

This new orientation in teaching [1], where related-instructional humour is infused and laced into lesson plans, could even be seen as essential to students population such as those in displaced and re-settled communities of South Sudan. These students have experienced severely disrupted socio-cultural-economic lives [35, 49] and are perceived to be more consumed by their day-to-day survival concerns rather than the learning of mathematics in the classroom setting.

Guiding this study is the theory of the teacher's communication competence in the classroom setting [17, 26, 37–41] which falls under constructivism. This theory argues that most creative, imaginative, and effective teachers always acquire and possess certain communication traits, qualities, or characteristics such as teacher's classroom immediacy, teacher's instructional quality, teacher's clarity in organisation skills, and teacher's socio-communication style in the classroom [37, 42, 43]. Among these effective teaching qualities, traits or characteristics is a teacher's humour orientation or the use of humour as a teaching and learning tool [31, 53]. When students view their teachers as using humour frequently, effectively, and appropriately, they also view them as more immediate, approachable, and friendly in the classroom. Hence, a teacher's perceived acceptance of the use of humour as a teaching and learning tool in the classroom is an indication of a teacher's overall communication competence and good teaching practices [45–47].

3. Relationship of the study to STEM education

Science, technology, engineering and mathematics (STEM) education is always described as an integrated-specialised way of teaching and learning with many various associated-acceptable definitions [48, 49]. However, the most preferable definition by STEM scholars is that STEM itself (known by the acronym science, technology, engineering, and mathematics) is a learning collaborative environment where students broaden their knowledge and learn through the process of exploration, invention, and discovery using real-world problems and situations [49]. The relationship of this study, however, about the use of humour in mathematics classroom is not only about the fact that mathematics as a subject matter itself is an embedded-integral part of STEM education as indicated by the STEM acronym in STEM literature, but also the fact that the use of humour or the application of the concept of humour in STEM education classroom settings would help connect the more specialised STEM education to other wider equally-valuable various knowledge domains out there, other non-STEM disciplines such as health, humanities, philosophy, psychology or creative arts to mention but few. The idea is that there may be other non-STEM alternative disciplines that might have been perhaps intentionally or unconsciously left out un-prioritised, underfunded, or undervalued, but that are equally contributing in various proportions in our attempt to understand or get a glimpse of the whole picture of our universe, or the multiverses as the saying goes. While STEM education is no doubt valuable in its own rights and efficient as it helps in quickly achieving the much desired economic prosperity, it is by itself alone not sufficient education, especially if promoted in an apparent expense or negligence of the other equally-valuable disciplines such as the creative fine arts, since those other forms of non-STEM education out there are also needed for the people's overall general well-being as fully developed human beings with an extra sense of humour. This is because having a sense of humour is documented in the literature as a sign of human strength, intelligence, wisdom and psychological maturity [5].

The use of humour as a teaching and learning technique in STEM education classroom settings is often rare if not almost non-existence as the humour use appears to be only implied or implicitly embedded in STEM's various educational robotic platforms, tools, or sources [49, 50], along with the associated educational robots, probably in the implicit forms of perhaps funny-humorous cartoons, interactive-computer games or collaborative group work. Therefore, there is still a need for explicit-appropriate humour use (such as the content-specific related-humorous stories, serious intellectual discussion or conversation, and even some civilised level of heated conceptual debates) from real sources or human beings combined or in conjunction with the use of educational robotic tools or sources in the STEM educational robotic platforms and STEM-related classroom settings. Literature review shows that the use of the concept of humour as teaching and learning tool in its various forms is more promoted, developed and therefore more advanced in the other non-STEM educational disciplines than it appears in the STEM education integrated subject areas such as mathematics classroom setting, where the use of humour (when present) is more often combative–rather than supportive–usually in the primitive forms of competitive riddles, numbers or wordplays [4]. The point or argument being expressed here is that all educational disciplines (whether STEM or non-STEM) are equal in terms of their contributions and values as they represent different dimensional-distinct abilities of various knowledge domains out there; and therefore focusing exclusively on just a few may only create certain levels of disabilities or deficiencies in others, deficiencies which can be minimised, supplemented or enhanced by the application of the concept of interdisciplinary integration, relationship or connection such as the use of humour (often associated with non-STEM education) in the STEM education classroom settings.

4. Methodology

Because the posed research question involved the exploration of views, attitudes, or opinions of teachers, the methodology found more suitable, appropriate, and adapted for this study was an opinion survey with close-ended questionnaire items or statements. About six (6) close-ended questionnaire items or statements (**Table 2**, first column) were prepared to survey the views, opinions, and attitudes of South Sudanese secondary school teachers about the use of humour in the classroom setting. A total of ten (10) secondary schools located within and around Juba city were surveyed, namely Juba Day secondary school, Juba Commercial, Nile Model, Rokon secondary, Supiri secondary, Rejaf secondary, Mat (Union), Hope and Future secondary schools. Sixty-five (65) South Sudanese secondary school teachers from ten (10) different secondary schools, three (3) of which were located in displaced and resettled communities, were asked to respond to the questionnaire items or statements shown in **Table 2**. This survey was in response to one of the volunteer teachers who first agreed to participate as co-teacher-researcher during the pilot study, but ended up opting out of the participation. This made the researchers to be curious and concerned about the general opinions or attitudes of South Sudanese secondary school teachers toward the use of humour in South Sudanese secondary school classrooms.

The questionnaire items or statements displayed in **Table 2** (first column) apply to the teaching and learning of mathematics because mathematics teachers are expected to be competent communicators in the subject matter, an acquired skill most mathematics teachers are not known to have mastered quite well. Mathematics teachers are often blamed and even accused, through teachers' evaluations or in the courts of public opinions, by many of their students [6, 7, 14, 45, 47]. They are often blamed for not always well explaining or even sometimes failing to satisfactorily explain even the lower level mathematics concepts in ways that are easily understood by the students-learners, lower-level concepts but important logical mathematics concepts such as, for example, why division by zero in never allowed, why negative times negative is always positive, why is any number raised to zero power is always one but never zero, or what exactly (if anything) is a zero raised to a zero power. Most of these allegations is due to mathematics teachers' widely perceived lack of communication competence in the subject matter [45], which is compounded by the perceived difficulties of teaching and learning the subject matter [47, 51]. A teacher's perceived competency in the use of appropriate types of humour [7, 31, 43, 54] in the classroom setting is an indication of a teacher's overall communication competence. Hence, the questionnaire items or statements displayed in the first column of **Table 2** apply to the teaching and learning of mathematics in the classroom setting.

The six (6) close-ended questionnaire items or statements were initially adapted from the literature-based Interpersonal Communication Competence Scales (ICCS) [17, 24, 26, 29, 37, 55–57] which were customised to form part of the Student Opinions Survey Questionnaires (SOSQ), a 35 questionnaire items-instrument with both close and open-ended questions or statements (parts A & B) intended to measure South Sudanese secondary school student's overall interest in the course materials [1, 10]. These items were developed to assess teacher's communication competence as one of the components-dimensions of interest in mathematics from the perspectives of the students-learners and were adapted here to also assess South Sudanese secondary school teachers' opinions, attitudes or views toward the use of humour as a teaching and learning tool. The close-ended section of SOSQ, part A, namely Adapted Literature-Based Interest Survey (Scale) Questionnaires (ALBISQ), from which the six (6) close-ended items or statements were adapted, had an overall alpha of 0.87, which shows a high degree of internal consistency of the items comprising that instrument. When tested for internal consistency (see **Tables 1** and **2**), the six (6) close-ended questionnaire items or statements showed an alpha of about 0.74, which is lower than the overall alpha of 0.87, but understandably so since the Cronbach's alpha coefficient is known to get lower as the number of items or statements gets fewer [17].

The views, opinions, or attitudes of South Sudanese secondary school teachers were captured using the above mentioned six (6) close-ended questionnaire items or statements (**Table 2**), and adapted Likert–like five-point scale instrument calibrated with semantic-differential statements and indicators such as *strongly negative* or *strongly positive,* etc., and was analysed according to SOSQ method of analysis [1, 10]. Each teacher's opinions responses which indicated either positive or negative views, opinions, or attitudes on each of the six (6) questionnaire items or statements were coded, organised, and quantified by assigning numerical values or codes to qualitative indicators as follows: Strongly Agree/Negative (1), Disagree/Negative (2), Neutral (3), Agree/Positive (4), and Strongly Agree/Positive (5). The resulting numerical

Cronbach's Alpha	No. of items
0.735	6

Table 1.
Reliability statistics.

	N	Minimum	Maximum	Mean	Std. deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. error	Statistic	Std. error
q1. The use of humour by the teacher helps reduce tension, anxiety, or stress	65	1.00	5.00	4.2615	0.95651	−1.991	0.297	4.640	0.586
q2. Students feel more at ease to participate in a class where humour is used	65	1.00	5.00	4.1231	0.83867	−1.551	0.297	3.439	0.586
q3. Humour helps hold students' attention and keeps the class interesting	65	1.00	5.00	4.1846	0.86408	−1.571	0.297	3.214	0.586
q4. The use of humour helps create a positive classroom environment	65	1.00	5.00	4.0923	0.80473	−1.470	0.297	3.623	0.586
q5. Humour should not be used to embarrass, ridicule or humiliate a student	65	1.00	5.00	4.0462	1.12404	−1.047	0.297	−0.039	0.586
q6. Students are more likely to attend a class where a teacher uses humour as a learning tool	65	1.00	5.00	4.0615	0.99808	−1.586	0.297	2.710	0.586
Valid N (listwise)	65								

Table 2.
Descriptive statistics.

values or codes for each of the six (6) questionnaire items or statements corresponding to teachers' responses were then arranged and organised in a way that allows descriptive patterns of views, opinions, or attitudes to emerge as generalised average rows and columns percentages (see **Table 4**, and its attached-explanation in the Appendix), somehow a continuous analogue of a discrete five-point Likert's like-continuum scale. The rows percentages indicate an individual teacher's self-reported scores while the columns' percentages show the group's average-self-reported scores corresponding to each of the six questionnaire items or statements (From here on, drop referring to Appendix while maintaining the **Table 4** referencing).

5. Results of the survey

Table 3 below, and drop Appendix referencing after (**Table 4**, Appendix) as average-columns-percentages responses corresponding to survey questionnaire items or statements (Q1-6 or S1-6) about the use of humour in the classroom as a teaching and learning tool. Interpreting and reading **Table 3**, for example, 86% of the average opinions of the sampled South Sudanese secondary school teachers agreed that the use of humour in the classroom helps reduce tension, anxiety, or stress (Q1, column 2). Similarly, 81% of the sampled teachers' average opinions agreed that humour should not be used to embarrass, ridicule, or humiliate a student (Q5, column 6) and so on. On this opinion scale continuum, 20% cut off would have been the lowest self-reported score indicative of the would be most the negative view recorded, 60% the average-neutral view or undecided responses (60% instead of 50% as this scale is skewed to the right by 10%) and 100% the highest possible maximal positive view achieved by some individual teachers: All the average-column-percentages in **Table 3** are all above the average neutral view (60%) and the overall average mean for all the columns is a favourable 83% average positive view or teachers' agreement toward the use of humour in the classroom setting as a teaching and learning tool.

We do not have results for the opinions of the individual teacher–the co-teacher-researcher–who opted out of the research participants and whom apparent quitting motivated this survey in the first place. We tried but could not reach him as he relocated to a different state during a follow-up interview designed to record his individual opinions, a follow-up which we could not do immediately upon his drop out because it would appear like coercion since his participation was voluntary. Although the posed research question explored the general views of South Sudanese secondary school teachers, the opinions of this particular-individual teacher would be interesting to know. All the individual opinions of the surveyed 65 secondary school teachers appear as rows of percentages in **Table 4**; and their overall general views relative to the six [7] questionnaire items or statements appear as columns percentages in **Table 4**, which are captured and summarised below as **Table 3**.

N = 65	Q1	Q2	Q3	Q4	Q5	Q6	Mean
Percentages (%)	86	82	84	82	81	81	83

Table 3.
Summary of the South Sudanese secondary school teachers' views toward the use of humour in the classroom as a teaching tool, technique, or strategy.

Teacher	Q1	Q2	Q3	Q4	Q5	Q6	Total	Percentages (%)
1	1/5	2/5	4/5	4/5	2/5	2/5	15/30	50
2	3/5	4/5	2/5	4/5	3/5	4/5	20/30	67
3	4/5	5/5	4/5	4/5	5/5	4/5	26/30	87
4	1/5	2/5	2/5	2/5	5/5	1/5	13/30	43
5	5/5	4/5	5/5	5/5	4/5	4/5	27/30	90
6	5/5	5/5	5/5	5/5	5/5	5/5	30/30	100
7	4/5	4/5	4/5	4/5	4/5	4/5	24/30	80
8	4/5	4/5	4/5	4/5	4/5	4/5	24/30	80
9	5/5	4/5	4/5	5/5	5/5	4/5	27/30	90
10	4/5	4/5	4/5	4/5	4/5	4/5	24/30	80
11	5/5	4/5	5/5	5/5	5/5	5/5	29/30	97
12	5/5	5/5	5/5	5/5	5/5	4/5	29/30	97
13	4/5	5/5	4/5	5/5	5/5	5/5	28/30	93
14	4/5	4/5	4/5	4/5	5/5	4/5	25/30	83
15	4/5	4/5	5/5	4/5	5/5	5/5	27/30	90
16	5/5	5/5	4/5	4/5	5/5	4/5	27/30	90
17	5/5	4/5	5/5	4/5	3/5	4/5	25/30	83
18	4/5	4/5	2/5	4/5	4/5	4/5	22/30	73
19	5/5	4/5	5/5	4/5	4/5	5/5	27/30	90
20	4/5	5/5	4/5	4/5	3/5	4/5	24/30	80
21	5/5	4/5	5/5	4/5	4/5	5/5	27/30	90
22	4/5	2/5	4/5	2/5	4/5	4/5	20/30	67
23	1/5	1/5	4/5	1/5	5/5	1/5	13/30	43
24	5/5	4/5	4/5	4/5	5/5	1/5	23/30	77
25	4/5	5/5	1/5	3/5	1/5	2/5	16/30	53
26	4/5	4/5	4/5	4/5	5/5	5/5	26/30	87
27	4/5	4/5	4/5	4/5	5/5	4/5	25/30	83
28	4/5	4/5	4/5	4/5	2/5	4/5	22/30	73
29	4/5	4/5	4/5	4/5	2/5	4/5	22/30	73
30	5/5	4/5	5/5	4/5	4/5	4/5	26/30	87
31	4/5	5/5	4/5	5/5	4/5	5/5	27/30	90
32	4/5	4/5	4/5	2/5	5/5	2/5	21/30	70
33	5/5	5/5	5/5	4/5	5/5	5/5	29/30	97
34	5/5	5/5	4/5	4/5	4/5	4/5	26/30	87
35	3/5	3/5	3/5	4/5	5/5	5/5	23/30	77
36	4/5	5/5	4/5	4/5	4/5	3/5	24/30	80
37	5/5	4/5	5/5	4/5	4/5	5/5	27/30	90
38	5/5	4/5	5/5	3/5	5/5	4/5	26/30	87

Teacher	Q1	Q2	Q3	Q4	Q5	Q6	Total	Percentages (%)
39	4/5	4/5	5/5	4/5	2/5	5/5	24/30	80
40	5/5	4/5	4/5	4/5	5/5	4/5	26/30	87
41	5/5	4/5	4/5	5/5	5/5	4/5	27/30	90
42	5/5	5/5	5/5	5/5	4/5	5/5	29/30	97
43	4/5	4/5	4/5	3/5	4/5	4/5	23/30	77
44	5/5	5/5	4/5	4/5	5/5	5/5	28/30	93
45	4/5	5/5	5/5	5/5	5/5	5/5	29/30	97
46	5/5	4/5	5/5	4/5	2/5	5/5	25/30	83
47	4/5	4/5	4/5	4/5	2/5	4/5	22/30	73
48	4/5	4/5	4/5	4/5	4/5	5/5	25/30	83
49	5/5	4/5	5/5	4/5	2/5	5/5	25/30	83
50	5/5	5/5	2/5	4/5	5/5	4/5	25/30	83
51	5/5	4/5	5/5	5/5	5/5	4/5	28/30	93
52	4/5	4/5	4/5	4/5	2/5	3/5	21/30	70
53	5/5	5/5	5/5	5/5	5/5	4/5	29/30	97
54	5/5	4/5	4/5	4/5	4/5	4/5	25/30	83
55	5/5	4/5	5/5	5/5	4/5	5/5	28/30	93
56	5/5	5/5	5/5	5/5	5/5	5/5	30/30	100
57	4/5	4/5	5/5	4/5	5/5	4/5	26/30	87
58	4/5	5/5	5/5	5/5	5/5	4/5	28/30	93
59	4/5	4/5	4/5	4/5	4/5	4/5	24/30	80
60	5/5	4/5	4/5	5/5	4/5	5/5	27/30	90
61	4/5	5/5	4/5	4/5	5/5	4/5	26/30	87
62	5/5	5/5	5/5	5/5	2/5	5/5	27/30	90
63	4/5	4/5	4/5	4/5	4/5	4/5	24/30	80
64	5/5	4/5	4/5	4/5	2/5	3/5	22/30	73
65	2/5	2/5	4/5	5/5	4/5	4/5	21/30	70
Total	278/325	267/325	272/325	266/325	263/325	264/325	1610/1950	5366/65
Percentages (%)	86	82	84	82	81	81	83	83

Table 4.
Views of South Sudanese secondary school teachers about the use of humour in the classroom as a pedagogical toolkit for teaching and learning mathematics in South Sudan.

6. Discussion of the results

The results above appear to indicate that the average majority of South Sudanese secondary school teachers are open and eager to experiment with new teaching tools or different types of teaching techniques and strategies such as the use of humour in teaching and learning. This keenness or eagerness to welcome new teaching and learning tools may be either because in South Sudan (being one of the newly emerging developing countries) people including teachers are hungry for education in order to catch up with the rest of the world [36, 58, 59]; or it could be due to the fact that classroom teachers have often been documented to always-generally show positive attitudes rather than negative opinions toward any research based-evidence, suggestions or new teaching and learning tools [13, 49]. This survey appears to imply that if offered any opportunities for professional growths or developments, the average majority of South Sudanese secondary school teachers may focus not only on *what* to teach (content-wise) but also explore the necessary-related pedagogy factors or dimensions such as the arts and science of *how* to teach creatively, imaginative, effectively, appropriately and reflectively in the classroom [19, 55, 56]. Teachers who tend to focus only on *what* to teach while ignoring or neglecting the *how* to teach aspects risk increasingly becoming perceived as just the *content persons* instead of being positively viewed as well-rounded professional educators. A *content person*, sometimes called a restricted or limited professional [60], is a common-low opinion jargon used in education circles or literature to describe teachers who either ignore, neglect, or just fear to explore the other necessary aspects of teaching (e.g., the *how* to teach factor) such as the teacher's creative yet effective pedagogical toolkit for teaching and learning a subject matter.

There are in general three types of teachers practicing in the classroom setting, namely the unprofessional, limited-restricted professional, and extended professional [60]. The extended professional is a continuously developed classroom teacher who regularly attends and actively participates in professional gatherings such as workshops, seminars, or academic conferences. In contrast, the unprofessional type of teacher is characterised by chronic absence from the work place, showing up to the class with unprepared or unrevised lessons, is often isolated from colleagues while at the same time hostile to students, and relies on the heavy use of corporal punishment as teaching techniques or strategy: These types of teachers tend to teach through threats, fear, and intimations as a teaching and learning strategy. The other third type of teachers, described as the limited-restricted professionals, are concerned mostly with the mastery of the content materials and/or skills often in the form of drills, repeated recitations, or rote memorisation techniques or strategies. These third types of teachers are either self-centered or concerned only with basic competence and tend to blame students for the failure to learn the materials. These types of teachers appear to have little or no continuous professional growth or development and are more often than not unimaginative. Hence, they are rigid as they appear to rely heavily on daily classroom routines as a form of teaching and learning strategy or technique in the classroom setting [60].

In contrast to the unprofessional or restricted types of teachers, the types of teachers known as extended professionals are the ones who would be expected to go beyond the technical competency of the subject matter. This is because these types of teachers would master not only *what* to teach but also *how* to teach effectively. They would take active responsibilities not only for themselves but also for their students-learners. In sum, these teachers are more often than not student-centered, adaptive, reflective, highly flexible, and independently or developmentally minded as well as creative thinkers [60]: These are the ones who would be expected to go the extra mile in terms of exploration, utilising or welcoming of humour as a possible pedagogical teaching and learning tool. This survey appears to show that the average majority of South Sudanese secondary school teachers have the potential to become student centered-practicing extended professionals as indicated by their positive attitudes, opinions, views, or beliefs toward the proposed and alternative new ways of teaching and learning in the classroom setting. South Sudanese secondary school teachers may have diverse academic, professional, and cultural backgrounds, but their converging average positive attitudes or views toward the use of humour in the classroom setting are encouraging.

7. Recap or wrap-up remarks

This opinions study surveyed the views of South Sudanese secondary school teachers toward the use of humour in the classroom as a teaching and learning tool. It was observed that the average majority of South Sudanese secondary school teachers are not only open but keen and eager (as suggested by their overwhelmingly positive average opinions) to welcome new ways of teaching and practicing in the classroom. The expressed views, attitudes, or opinions (referring to **Tables 3** and **4**) are indicative of positive disposition toward the use of appropriate types of humour in the classroom setting. Hence, a newly proposed-alternative and equivalent method of teaching mathematics such as H-SIA appeared to be welcomed by the average majority of South Sudanese secondary school teachers. However, the extent to which the South Sudanese secondary school teachers are able to put the disposition (humour) to practice in the classroom setting is not yet known.

This opinion survey was, however, limited only to 65 South Sudanese secondary schools teachers from ten (10) secondary schools located within and around Juba city. Three out of ten (3/10) of these surveyed schools were located in displaced and re-settled communities. The survey would have been more convincing had it included representative samples from all the former ten states of South Sudan, which were then momentarily inflated into more than thirty-two (32) politically motivated-controversial states before their reinstatement back last year into the original ten (10) states plus three (3) more administrative areas. There are approximately about 70 secondary schools located within and around Juba city, with fifteen (15) of them being public or government-run schools while the rest of these schools are privately sponsored. The next study could take a look at a larger randomised sample of South Sudanese secondary school teachers and classify the teachers' views, opinions, or attitudes by the corresponding subject matters. Further study can also attempt to identify what percentage (if any) of South Sudanese secondary school teachers actually use humour in their teaching practices.

It is then recommended that classroom teachers, particularly mathematics teachers, be encouraged and always allowed a certain degree of freedom to explore and experiment with new ways of teaching and learning such as the use of humour as a teaching and learning tool. However, in order for this to be implemented effectively, it is suggested that classroom teachers be first properly trained on how to use humour appropriately, creatively, and effectively in the classroom setting.

Acknowledgements

The authors are grateful to the South Sudanese secondary school teachers who positively responded and voluntarily participated in this secondary school teachers' opinion survey.

Appendix

Table 4 shows the details of the data analysis about how the teachers' self-reported Likert's styled scores (reflecting teachers' views, opinions, or attitudes) were arranged by questionnaire items or statements (e.g., Q1-6 or S1-6), the quantified teachers' views arranged in a rectangular array spreadsheet–excel like format, and then organised into descriptive patterns of generalised-iterated rows and columns sums average percentages (Tap et al., 2019, 2020), where the sums of the rows show individual teacher's average percentage on each of the six questionnaire items or statements. Meanwhile, the sums of the columns show groups' or teachers' average percentages on each of the six items or statements (Q1-6). On this somehow continuous analogue of a discrete Likert's five-point scale of South Sudanese teachers' opinions scale (SSTOS), the lowest opinion (which was never expressed or recorded) would have been at 20% (65/325) cut off, the medium cut off, neutral or undecided opinion was at 60% (195/325) cut off and the highest possible opinion (which was achieved by some individual teachers or outliers) was at 100% (325/325). To make the Tables more readable, the average percentages were rounded to the nearest whole numbers.

Author details

William Deng Tap*, Helicopter Mark Bulbul and Biar Simon Ajang
School of Education, Department of Mathematics, University of Juba, Juba,
South Sudan

*Address all correspondence to: tapx0001@umn.edu

References

[1] Tap WD. Relative effects of a humour-supported instructional approach on interest in mathematics [among secondary school students in South Sudan re-settled communities (Unpublished Doctoral Dissertation)]. Harare: University of Zimbabwe; 2020. p. 2020

[2] Zazkis R, Liljedahl P. Teaching mathematics as storytelling. Rotterdam/Taipei: Sense Publishers; 2009

[3] Torok SE, McMorris RF, Lin WC. Is humor an appreciated teaching tool? Perceptions of Professors' teaching styles and use of humor. College Teaching. 2004;**52**:14-20

[4] Grawe PH. A meeting of minds: An alternate humor for teaching mathematics to non-stem majors. Journal of Humanistic Mathematics. 2016;**6**:220-234

[5] Oosthuizen RM. Resilience as moderator between workplace humour and well-being, a positive pyschology perspective. In: Vanderheiden E, Mayer CH, editors. The Palgrave Handbook of Humour Research. Switzerland, Cham: Spring Nature, Palgrave Macmillan; 2021. pp. 263-287

[6] Deiter R. The use of Humor as a Teaching Tool in the College Classroom. 2015. Available from: https://www.merlot.org/merlot/viewMaterial.html?id=329565 [Accessed: October 15, 2015]

[7] Steele KE. The positive and negative effects of humor in the classroom setting (ERIC Document Reproduction Service No. ED426929). s.l. : ERIC, 1998. Dissertations/Theses-Master Theses (042); 1998

[8] St-Amand J, Smith J, Beland S, Moreau D. Understanding teacher's humor and its attributes in classroom management: A conceptual study. Educational Sciences: Theory and Practice. 2021;**21**(2):115-130

[9] Garner RL. Humor in pedagogy: How ha-ha can lead to aha! College Teaching. 2006;**54**:177-180

[10] Tap WD, Mtetwa DK, Vere JC. Using humor for generating and maintaing Interest in mathematics among secondary school students in South Sudan Re-settled communities. African Journal of Research in Mathematics, Science and Technology Education. 2019;**23**:244-253

[11] Struthers J. The case for mixed methodologies in researching the teacher's use of humour in adult education. Journal of Further and Higher Education. 2011;**35**:439-459

[12] Repass JT. Making test anxiety a laughing matter: A quantitative study (Unpublished doctoral dissertation, University of Keiser, ProQuest 10261256). 2017. Using humour for generating and maintaining interest in mathematics among secondary school students in South Sudan re-settled communities; 2017

[13] Elkhayma R. Attitudes toward the use of humor: A case study of Moroccan secondary school teachers. In: Proceeding of IASTEM International Conference. Rabat: IASTEM; 2021

[14] Godbey C. Mathematics anxiety and the underprepared student (ERIC Document Reproduction Service No. ED426743). s.l. : ERIC, 1997. Guide-Classroom-Teacher (052); 1997.

[15] Durik AM, Matarazzo KL, Delayey ML. The effects of humorous instructional materials on interest in

mathematics task. Motivation and Emotion. 2010;**343**:293-305

[16] Tews MJ, Jackson K. Fun in the college classroom: Examining its nature and relationship with student engagement. College Teaching. 2015;**63**:16-25

[17] Kromka SM, Goodboy AK. Classroom storytelling: Using instructor narratives to increase student recall, affect, and attention. Communication Education. 2019;**68**:20-43

[18] Green MC. Storytelling in teaching. Pyschologicalscience.org. 2015. Available from: http://www.psychologicalscience.org/observer/getArticle.cfm?id=1562 [Accessed: September 5, 2015]

[19] Maguire J. Creative Storytelling: Choosing, Inventing, and Sharing Tales for Children. New York: The Philip Lief Group; 1985

[20] Eggleton PJ. Motivation: A key to effective teaching. Mathematics Educator, Two-Year College Mathematics Journal. 1992;**3**:1-12

[21] Schukajlow, S (2015.) Effects of enjoyment and boredom on students' interest in mathematics and vice-versa. Hobart, Australia: PME, 2015. Proceedings of 39th Psychology of Mathematics Education (PME) Conference. pp. 137-144.

[22] Weber K. Mathematical humor: Jokes that reveal how we think about mathematics and why we enjoy it. Mathematical Intelligencer. 2016;**38**:56-61

[23] Masek A, Hashim S, Ismail A. Integration of the humor approach with student's engagement in teaching and learning sessions. Journal of Education for Teaching. 2018;**45**:228-233

[24] Wanzer MB, Frymier AB. The relationship between student perceptions of instructor humor and students' reports of learning. Communication Education. 1999;**48**:48-62

[25] Warwick J. An experiment relating humor to student attainment in mathematics. PRIMUS. 2009;**19**:329-345

[26] Karlin NJ, Machleu M. The relationship between instructor use of different types of humor and student interest in the course material. College Teaching. 2017;**65**:192-200

[27] Goodman-Scott E. Enhancing student learning by "Building Caring Climate": School counellors' experiences with classroom management. Professional School Counselling. 2019;**22**:1-12

[28] Gadanidis G, Gadanidis JM, Huang Y. Using humor to gain mathematical insights. National Council of Teachers of Mathematics (NCTM). 2005;**10**:244-250

[29] Tap WD, Mtetwa DK, Vere JC. Types of humour categories used to generate and maintain interest in mathematics among secondary school students in South Sudan's displaced and re-settled communities. Journal of Education and Training Studies (JETS). 2021;**9**:14-23

[30] James DL. Split a gut and learn: Theory and research (ERIC Document Reproduction Service No. ED458671). s.l.: ERIC, 2001. Information analysis (070); 2001

[31] Weimer M. Humor in the classroom: 40 years of research. Faculty Focus. 2013. Available from: http:www.facultyfocus.com/articles/effective-teaching-strategies/humor-in-the-classroom-40-years-of-research [Accessed: September 5, 2016]

[32] McGraw P., Warner J. What, Exactly, Makes Something Funny? A Bold New Attempt at a Unified Theory

of Comedy. Slate.com. 2001. Available from: http:www.slate.com/articles/arts/culturesbox/features/2014 [Accessed: September 9, 2019]

[33] Morreall J. Humor, philosophy and education. Educational Psychology and Theory. 2014;**46**:120-131

[34] McNeely R. Using humor in the classroom: Laughter has the power to fuel engagement and help students learn. National Education Association. 2019. Available from: http:www.nea.org/tools/52165.htm [Accessed: October 15, 2019]

[35] Nishad TM. Mathematical theology—A new approach. International Journal of Mathematics Trends and Technology. 2021;**67**:68-70

[36] Kuek JC et al. Hunger for an Education: A Research Essay on the Case of South Sudan and the Voice of Its People. Forum for International Research in Education (FIRE). 2014;**1**(2):1-12. DOI: 10.18275/fire201401021004

[37] Anderson JF. Teacher immediacy as a predictor of teaching effectiveness. [ed.] B. D. Ruben. 1, 1979. Communication Yearbook. 1979;**3**:534-559

[38] Spitzberg BH. Communication competence as knowledge, skill and impression. Communication Education. 1983;**32**:323-329

[39] Rubin RB, Martins MM. Development of a measure of interpersonal communication Competence. Communication Research Reports. 1994;**11**:33-44

[40] Richmond VP. Communication in the classroom: Power and motivation. Communication Education. 1990;**39**: 181-195

[41] Hartley P. Interpersonal communication. 2nd ed. London and New York: Routledge; 1999. p. 1999

[42] Anders SL, Tucker JS. Adult attachment style, interpersonal communication competence, and social support. Personal Relationships. 2000;7:379-389

[43] Wanzer MB, Frymier AB, Irwin J. An explanation of the relationship between instructor humor and student learning: Instructional humor processing theory. Communication Education. 2010;**59**:1-18

[44] Kane J. Education, information, and transformation. 6th ed. Prentice-Hill: Upper Saddle River, NJ; 1999

[45] Ornstein AC, Behar-Horenstein LS, Pajak EF. Contemporary Issues in Curriculum. 3rd ed. Boston: Allyn and Bacon; 2003. p. 2003

[46] Driscoll MP. Pyschology of Learning. 3rd ed. Boston: Pearson; 2003. p. 2005

[47] Nyaumwe L. Professional development of secondary school mathematics teachers through collaborative reflection in pre-service and in-service learning context (Unpublished doctoral dissertation, University of Zimbabwe); 2008

[48] Chatzopoulos A, et al. Innovative robot for educational robotics and STEM. In: Kumar V, Troussas C, editors. Intelligent Tutoring Systems: Lectures Notes in Computer Science. Vol. 2020. Switzerland, Cham: Springer; 2020. pp. 95-104

[49] Vlasopoulow M, Kalogiannakis M, Sifaki E. Investigating teachers' attitudes and behavioral intentions for the impending integration of STEM education in primary school. In: Papadakis S, Kalogiannakis M,

editors. Handbook of Research on Using Education Robotics to Facilitate Student Learning. Vol. 2021. Hershey, PA: IGI Global; 2021. pp. 235-256

[50] Dorouka P, Papadakis S, Kalogiannakis M. Tablets and apps for promoting robotics, mathematics, STEM education and literacy in early childhood education. International Journal of Mobile Learning and Organisation. 2020;**14**:255-274

[51] Papo WD. Effective teaching and learning in large classes at tertiary institutions (Unpublished doctoral dissertation, Rand Afrikaans University); 1997

[52] Fitriah F. Appropriate and inappropriate uses of humor by teachers and the effects of it in learning. Linguistics, Literature, and Language Teaching (LET). 2012;**2**:1-7

[53] Banas J, et al. A review of humor in educational settings: Four decades of research. Communication Education. 2011;**60**:115-144

[54] Winsor MS, Kirwan JV, Ssebaggala L. What do we know about secondary mathematics teacher preparation in the United States? The Mathematics Educator. 2018;**27**:73-106

[55] Wardman NP. umour or humiliation? When classroom banter becomes irresponsible sledging in upper-primary school contexts. Discourse: Studies in the Cultural Politics of Education. 2020;**1**:1-12

[56] Puggina AC, Silva MJ. Interpersonal communication competence scale: Brazilian translation, validation and cultural adaptation. Acta Paul Enferm. 2014;**27**:108-114

[57] Palacious A, Arias V, Arias B. Attitudes toward mathematics: Construction and validation of a measurement instrument. Revista de Psicodidactica. 2014;**19**:76-91

[58] Masungu KA. Delphi Survey Process on a Decentralized Education System in Southern Sudan (Dissertation Abstract International). Dissertation Abstract International; 2012

[59] Brown G. Education in Southern Sudan: Investing in a Better Future. London, England: Centre for Universal Education, Brooking Institute; 2011. p. 2011

[60] Davies L, Harber C, Schweisfurth M. Democratic Professional Development. Birmingham: CIER/CFBT; 2005

Chapter 4

Impact of Integrated Science and Mathematics Instruction on Middle School Science and Mathematics Achievement

Zenaida Aguirre-Muñoz, Brian Dang and Elias S. Loria Garro

Abstract

Despite the impetus from professional organizations for science and math integration, evidence in support of such efforts in raising both science and mathematics achievement is scarce, particularly for underrepresented students. The available literature is mixed especially regarding impact on mathematics outcomes. This exploratory study documents the impact of the Middle School Math and Science $(MS)^2$ Integration project based on the results of the internal evaluation of an intensive teacher training model for integrated science and mathematics in middle school. Multivariate analysis of variance shows $(MS)^2$ positively impacted middle school students' science and mathematics knowledge in this sample of diverse students. Overall, the $(MS)^2$ group outperformed the comparison group. There was also evidence that students who received $(MS)^2$ integrated instruction in science classrooms slightly outperformed those who received $(MS)^2$ integrated instruction in math classrooms. Multiple regression results indicated that $(MS)^2$ group membership and opportunity to learn through integrated instruction were significant predictors of students' science and mathematics scores. Although students in $(MS)^2$ classrooms were more likely to have higher achievement scores, the frequency of integrated instruction opportunities also significantly predicted student achievement, particularly in mathematics classrooms. Ethnicity and gender were not significant predictors of student scores. Implications are discussed.

Keywords: integrated science and mathematics instruction, mathematics achievement, middle school, science achievement

1. Introduction

Science and mathematics integration has become increasingly popular in recent years. In the USA, the National Council of Teachers of Mathematics [1] and the National Science Teachers Association (NSTA) [2], continuously highlight the benefits of integrating science and mathematics to improve the quality of education. Education researchers have advocated for the implementation of educational initiatives that emphasize a science and mathematics integrated curriculum, such as the Mathematics Common Core State Standards (CCSS-M) and the Next Generation Science Standards (NGSS) [3]. A well-recognized benefit of integrated instruction is its alignment with student-centered approaches to learning that promote critical thinking (e.g., [4–6] interest [7], motivation [8] and cognitive engagement [9–11]. Indeed, integrated STEM (science, technology, engineering, and mathematics) can shift and shape STEM beliefs and engagement [12]. Despite the promise of these efforts, additional evidence of its impact on student science and math achievement is needed [13, 14].

Thus, the main purpose of this study was to examine the impact of integrated science and mathematics instruction on middle school student achievement and to examine if the impact varies by classroom disciplinary context (i.e., whether integrated instruction occurred in mathematics classrooms or science classrooms) or by student gender. A secondary purpose is to examine the extent to which integrated science and mathematics instruction is a predictor of student science and mathematics achievement beyond student ethnicity, gender, or classroom disciplinary context. These analyses will also provide information about the degree to which the impact of integrated instruction is shared equally among student of different ethnicities and genders.

1.1 What is integrated instruction?

Although there are several definitions of science and mathematics integration, three definitions serve as the foundation for the definition applied in this study. Ref. [15] proposed the Berlin-White Integrated Science and Mathematics (BWISM) model that emphasizes the need to integrate big ideas that are common in both science and mathematics to enhance integration of the curriculum. A second definition, proposed by [16] underscores integrated instruction happens when there is alignment between content of the CCSS-M and NGSS. For instance, two disciplinary practices that align with the CCSS-M and NGSS are the use of arguments. The NGSS promotes activities in which students engage in arguments from evidence, while the CCSS-M encourages students to make viable arguments and critique the reasoning of others when learning new mathematical concepts [13]. Therefore, teachers can design activities in which students make interdisciplinary connections to further understand specific topics pertaining to both science and mathematics. These two integration perspectives draw attention to the synergistic relationship between science and mathematics content and disciplinary practices. These synergistic relationships refer to concepts that mutually correspond to both science and mathematics, which may also overlap in disciplinary norms, cognitive demand, or complex problem-solving [13]. For instance, teachers can integrate science and math instruction by teaching students inductive and deductive reasoning strategies to find patterns in both science and mathematics activities [15]. In this context, teachers exploit mathematical concepts such as differential equations (functions) to teach about exponential growth and decay within ecosystems in relation to specific populations.

Consequently, instruction that emphasizes synergistic relationships between science and mathematics provides students with interdisciplinary connections that helps them understand and make deeper connections with content concepts [17–19]. The third definition proposed by [20, 21] highlights the importance of using the integration continuum model to assess progress towards fully integrated practice. This model reflects overlapping elements between mathematics and science in instruction. That is, movement along the continuum represents different levels of integration whereas the ends of the continuum signal non-integrated practice [21]. Informed by these approaches to integration, we define integration as instruction that "coordinates the teaching of science and mathematics through engagement with disciplinary practices and/or coordinating concepts" [13].

1.2 Impact of math science integration

Advocates for integrated instruction argue that (a) deep understanding depends upon connections between ideas and (b) integrated instruction better resembles problems people will face in the real-world [22]. Previous studies provide the theoretical justification as well as initial empirical evidence that integrated instruction leads to increases in math and science achievement [23, 24], student interest [25, 26], and creativity [27]. More recent quasi-experimental methods and randomized controlled trials provide further empirical evidence that the use of integration increases student interest [28], math and science achievement [29], and problem-solving skills [30]. For example, [30] separated 4th grade students into an integrated math-science group and a control group (N = 117). Both groups of students studied the same topics for 8 weeks and took a problem-solving skills pre-test and a post-test. Results indicated that the integrated instruction group had a larger increase in problem solving skills than the control group. Similarly, [29] collected science assessments and attitude surveys from students in 8th grade math-science integrated classrooms and compared them to non-integrated classrooms (N = 1695). They found students in integrated classrooms had higher confidence in graphing, understanding of math-science integrated concepts and female students outperformed male students in science. We build on these studies by examining the impact of integrated instruction on science and mathematics achievement and explore the relationship between opportunity to engage in integrated instruction and achievement. We also examine if these opportunities had a differential impact based on student ethnicity or gender.

1.3 Context for the study

This study was part of a larger research study investigating the effectiveness of an intensive middle school integrated science and mathematics teacher development project: The Middle School Math and Science Integration program $(MS)^2$. This larger project was a three-year state-wide teacher development program providing teachers with a master's degree in multidisciplinary science and mathematics instruction in a hybrid environment. Program coursework and activities were designed to develop teachers' content and pedagogical knowledge to implement integrated instruction. Teachers took synchronous online courses over the fall and spring semesters and engaged in intensive face-to-face instruction over extended summer sessions. Initial evidence of positive impact on teacher learning and practice has been published elsewhere [13]. The focus of this paper is to report on the impact of integrated instruction on middle school students' science and mathematics achievement.

The main goal of the larger teacher-training project was to increase teachers' pedagogical content knowledge (PCK) [31] Three education courses aimed to: (a) expose teachers multiple integration approaches; (b) provide multiple models of fully integrated lessons; and (c) provide meaningful opportunities to situate these lessons in culturally and linguistically diverse classroom contexts. Two mathematics and seven science content courses served to engage teachers in complex science and math content and problem solving intended to support teachers' development of content knowledge needed to construct and implement effective integration lessons. Three of the content science courses were taught in an integrated manner. In these courses, a science professor (i.e., biology or physics professor) co-taught with a mathematics professor to provide teachers with experience in integrated instruction at the professional level. Thus, were simultaneously enrolled in both a science and mathematics course that were designed to complement the content of the other discipline.

To support learning of science and mathematics disciplinary practices, teachers were also provided with opportunities to engage in mathematics and scientific communities through local and national research lab participation. Required research lab experiences further supported the development of these *habits of mind* (e.g., healthy skepticism, curiosity, appreciation for new ideas, etc.) and grounded their pedagogical content knowledge (PCK). If teachers are to create authentic communities of practice in their science and math classrooms, they need direct experiences in discipline-based communities of practice. These experiences have been linked to improved science teaching quality and improved student performance [32]. In their second year of program participation, teachers were invited to participate in research either locally or in national research labs. Finally, teachers collaborated with other teachers and STEM faculty to develop artifacts they used in their classrooms. Development of these materials also supported development of PCK contextualized to their classroom environments. A sample of teachers from this larger project was recruited to participate in the current study. We report on the impact of integrated science and mathematics instruction on student learning. The specific research questions addressed here include:

1. What is the impact of math and science integrated instruction on student achievement? Does math or science achievement vary by disciplinary context?

2. What classroom (disciplinary context and level of integration opportunities) and student factors (ethnicity, gender) significantly predict science and mathematics achievement?

2. Methodology

A 2 [instruction group: [$(MS)^2$ or comparison] by 2 (discipline context: math or science classroom) by 2 (gender: male or female) factorial research design was implemented to answer the research questions. **Table 1** presents the research design and sample sizes for each cell. For the first research question, a multivariate analysis of variance (MANOVA) with instruction group, discipline, and gender as fixed factors and

	$(MS)^2$		$(MS)^2$	Control		Control
	Male	***Female***	**Total**	***Male***	***Female***	**Total**
Math Class	66	79	145	81	56	137
Science Class	52	68	120	76	66	142
Total	118	147	265	157	122	279

Table 1.
Design and cell sample sizes.

mathematics and science achievement scores as dependent variables. For the second research question, two multiple regression analyses were performed with five predictor variables (instruction group, discipline context, integration opportunities, ethnicity, and gender) and mathematics or science achievement scores were the outcome variable.

2.1 Participants

2.1.1 Students

A total of 544 students [265 $(MS)^2$ and 279 comparison] representing 19 diverse classrooms [9 $(MS)^2$ and 10 comparison] from 19 schools in ethnically diverse schools throughout a state in the southwest of the United States participated in the study. Of these students, 50.5% (275) are male and 49.4% (269) are female. Student ethnicity/race was reported as follows: 36.2% (197) Latinx; 27.5% (150) African American; 25.7% (140) White; and 10.5% (57) other. Students were also roughly evenly distributed between math and science classes with 51.8% enrolled in a mathematics class and 48.2% enrolled in a science class. The great majority of students (97.6%) received free and reduced lunch. The comparison group was matched based on socio-economic status.

2.1.2 Teachers

Nineteen [9 $(MS)^2$, 10 comparison] middle school mathematics [9: 4 $(MS)^2$, 5 comparison] and science [10: 5 $(MS)^2$, 5 comparison] teachers participated in the study. One of the $(MS)^2$ teachers moved to another state in the middle of the year and thus was dropped from the study. Comparison group teachers were identified based on matching on three key teacher characteristics: years teaching in the discipline (math or science), gender (male or female), and completion of a master's degree related to the discipline they taught or an education-related degree (e.g., MA in Curriculum and Instruction). At the time demographic information was collected, about 32% of these teachers [6: 3 $(MS)^2$ and 3 comparison teachers] regularly taught two or more middle school grades, all taught 6th for at least one class period. Years of teaching experience ranged from 2 to 15, with an average of 7.23 years for the $(MS)^2$ teacher and 6.67 for the comparison teachers. All participants [$(MS)^2$ and comparison] hold certification in the content they taught. Teachers taught in schools where at least one third of the school student population is economically disadvantaged as indicated by state designation criteria (i.e., percent receiving free or reduced lunch). In addition, all the teachers at the start of the program taught in stand-alone mathematics or science classrooms. $(MS)^2$ teachers enrolled in and completed the program (MS in Science Education) because of their interest in integration. $(MS)^2$ teachers were asked to implement integrated instruction in their most ethnically diverse class. Teachers were identified based on having at least one class that had at least 50% students of color enrolled.

2.2 Instrumentation

2.2.1 Iowa test of basic skills (ITBS)

To determine the impact of integrated science and mathematics instruction, the Iowa Test of Basic Skills (ITBS)-are multiple-choice, standardized achievement tests for students in kindergarten through eighth grade- was administered to all students. The science and mathematics batteries were administered to students in April.

Students were given 60 minutes to complete the mathematics test and 60 minutes to complete the science test.

2.2.1.1 ITBS math

The math survey battery (63 items total) is comprised of three sections: (1) concepts and estimation section; (2) problem solving and data interpretation; (3) computation. The concepts and estimation section is comprised of 27 items that measure concepts such as numeration, properties of number systems, and number sequences; fundamental algebraic concepts; and basic measurement and geometric concepts- probability and statistics. The 20, estimation knowledge and skills items measure (a) standard rounding—rounding to the closest power of 10 or, in the case of mixed numbers, to the closest whole numbers; (b) order of magnitude involving powers of 10; and (c) number sense, including compatible numbers and situations that require compensation. The problem solving and data interpretation includes 6 items that measure "problem-solving process" or "strategy." Such items measure steps of (1) getting to know the problem, (2) choosing what to do, (3) doing it, and (4) looking back. The data interpretation skills assessed in this test are reading amounts, comparing quantities, and interpreting relationships and trends in graphs and tables. The computation section includes 10 items that assess a single operation—addition, subtraction, multiplication, or division on whole numbers, fractions, and decimals.

2.2.1.2 ITBS science

The science test consisted of 60 items from two ITBS forms to target more of the 6th grade curriculum. Items targeted content and process pertaining to four domains: scientific inquiry, life science, earth and space science and physical science. Scientific inquiry targets understanding methods of scientific inquiry and process skills used in scientific investigations. Life science assesses knowledge of characteristics of life processes in plants and animals; body processes, disease, and nutrition; continuity of life; and environmental interactions and adaptations. Earth and space science assesses the Earth's surfaces, forces of nature, conservation and renewable resources, atmosphere and weather, and the universe. Physical science targets basic understanding of mechanics, forces, and motion; forms of energy; electricity and magnetism; properties and changes of matter.

2.2.1.3 ITBS reliability and validity

Several dozen validity studies indicate the ITBS instruments have strong technical quality (e.g., [33, 34]) with K-R20 coefficients ranging from 0.87 to 0.933 for the mathematics and 0.90 for science batteries. Past studies also show high correlations with state assessments, 0.76–0.81 (e.g., [35]) and measures (e.g., [36]) as is predictive of grade point average (e.g., [37]).

2.2.2 Integrated science and mathematics opportunity to learn student (OTL) survey

Students were asked to report on the frequency of integrated instructional opportunities. The integrated instruction OTL survey was comprised of eight items asking how often content concepts represented integrated topics, how often the teacher explained how concepts were integrated, how often the teacher provided feedback prompting making connections between content areas and how often the lessons involved science inquiry or project-based learning. Students reported their response on a five-point scale ranging from "1" representing never or hardly ever, to "5" representing "every day." Internal consistency for this survey was .88, indicating strong reliability. Internal consistency on these items was strong (.92, Cronbach alpha).

2.3 Procedures

2.3.1 Teacher training

$(MS)^2$ teachers were trained on math and science integration as part of a hybrid master's program [13]. Teachers participated in a two-year state-wide program that provided teachers with a master's degree in multidisciplinary science and mathematics instruction in a hybrid environment. Program coursework and activities were designed to develop teachers' PCK to implement integrated instruction. Teachers took synchronous online courses during the fall and spring semesters and engaged in intensive face-to-face instruction during two summer sessions. A total of 12 courses included three education courses noted in the Context for the Study section above. Two mathematics courses targeted middle school mathematics and were integrated with biology content. Seven science content courses engaged teachers in complex science problem solving. These content courses supported teachers' growing development of content knowledge needed to construct and implement effective integration lessons. Six of the nine content courses were taught in an integrated manner with mathematics and science purposes of teaching integrated topics.

2.3.2 Lesson development

In each of the content courses, teachers developed integrated lesson plans in learning circles. Each learning circle consisted of one mathematics middle school teacher, one middle school science teacher and one faculty member. These triads developed integrated lesson plans in each of the content courses and the faculty member provided feedback on the integrity of the content concepts targeted by the lessons and the appropriateness for math and science integration. We have reported elsewhere that initially these lessons did not represent fully integrated content, but as they progressed in the program, the level of integration reflected synergistic concepts and improved student feedback [13]. Teachers used the lessons they developed during the master's coursework throughout the yearlong implementation of this study.

Comparison group teachers received 6 hours of regular district training in science inquiry and problem-based learning and attended the mandatory district training throughout the year of the study. Teachers administered the science and mathematics integration opportunity survey as well as the mathematics and science assessments in May of the year the study took place.

3. Results

3.1 Impact on student learning

The MANOVA results indicated there were statistically significant differences in achievement based on instructional group, F(2, 535) = 1004.48, $p < 0.001$; Wilk's

$\Lambda = 0.210$, partial $\eta^2 = 0.790$. Prior to examining the between-subject effects, the homogeneity of variance assumption was tested for both achievement measures. The Levene's F tests were statistically significant ($p < 0.05$), suggesting that the variances associated with math and science achievement were not homogenous. However, an examination of the standard deviations revealed that none of the largest standard deviations were more than four times the size of the corresponding smallest standard deviation (**Table 2**), indicating that the between-subject effects would be robust in this case [38].

Between subject effects showed the instructional group had a significant effect on both math and science achievement, $F(1, 536) = 1057.19$, $p < 0.001$; partial $\eta^2 = 0.663$ and $F(1, 536) = 1516.57$, $p < 0.001$, partial $\eta^2 = 0.739$, for math and science achievement, respectively. Math and science mean comparisons show the $(MS)^2$ group outperformed the comparison group regardless of gender and disciplinary context (**Figure 1**). Effect sizes (partial η^2) are large indicating the instruction group effect accounts for a large proportion of the variance in student scores. Disciplinary context did not significantly affect students' mathematics ITBS scores ($p = 0.229$). The disciplinary context main effect for the science ITBS scores, however, was nearly significant ($p = 0.051$). Given the observed power was below optimal levels (0.424), this result may be attributable to sample size [39]. Mean score differences in science achievement between students in math and science classes (**Figure 1**) was 0.12 which is not meaningfully different. Gender differences were also not significant (p's > 0.05). Thus, student math and science performance was comparable whether they were in math or science classrooms or whether they were male or female students.

Math Achievement								
Variable	**Math Class**				**Science Class**			
	$(MS)^2$ Group		**Control Group**		**$(MS)^2$ Group**		**Control Group**	
	x	***SD***	***x***	***SD***	***x***	***SD***	***x***	***SD***
Male	46.95	9.35	26.37	3.13	45.63	8.93	26.80	3.03
Female	46.46	11.33	26.13	2.90[1]	49.81	9.94	26.73	3.13
Total	46.68	10.44	26.27	2.03	48.00	9.708	26.77	3.07
Science Achievement								
Variable	**Math Class**				**Science Class**			
	$(MS)^2$ Group		**Control Group**		**$(MS)^2$ Group**		**Control Group**	
	x	***SD***	***x***	***SD***	***x***	***SD***	***x***	***SD***
Male	59.00	5.69	47.01	2.65[2]	61.37	1.59	46.17	4.30
Female	59.87	4.48	47.54	4.17	61.60	1.90	47.00	5.06
Total	59.48	5.07	47.23	3.35	61.50	1.77	46.56	4.67

[1] *The assumption that the largest standard deviation is not more than four times than the size of the corresponding smallest standard deviation is met [38]. In this case, 2.90 is the smallest SD and 11.33 is the largest SD. Since (2.90 × 4) = 11.6 and 11.6 > 11.33 the assumption is met.*

[2] *The assumption that the largest standard deviation is not more than four times than the size of the corresponding smallest standard deviation is also met. In this case, 2.65 is the smallest SD and 5.69 is the largest SD. Since (2.65 × 4) = 10.6 and 10.6 > 5.69 the assumption is met.*

Table 2.
Means and standard deviations for math and science achievement.

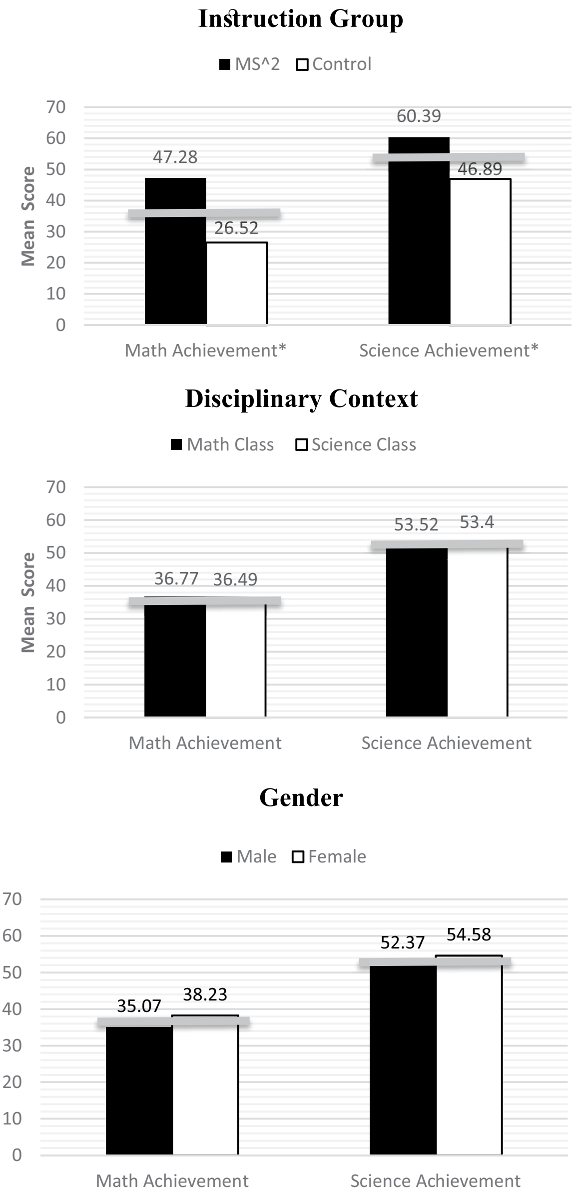
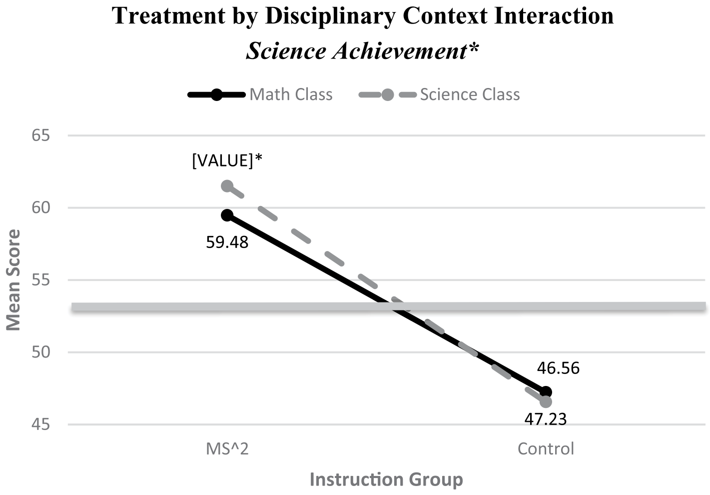


Figure 1.
*Instruction group, disciplinary context, and gender comparisons for mathematics and science achievement. Note: *statistically significant at the 0.001 level; gray line represents the grand mean.*

For math achievement, the instruction group by disciplinary context interaction was not significant, F(1, 536) = 0.154, p = 0.695, partial $\eta^2 < 0.001$. However, for science achievement, the instructional group by disciplinary context interaction effect had a significant effect on science achievement, F(1, 536) = 15.502, $p < 0.001$, partial $\eta^2 = 0.028$. Cell mean comparisons show the $(MS)^2$ group performance advantage depended slightly on disciplinary context (**Figure 2**). That is, $(MS)^2$ students in science classes performed slightly higher on the science ITBS (x = 61.50) than $(MS)^2$



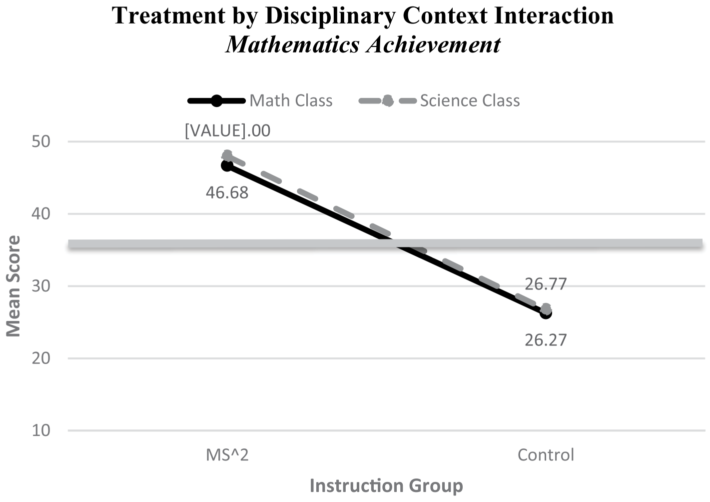


Figure 2.
*Interaction effects for science and mathematics achievement. Note: * statistically significant at the 0.001 level; gray line represents the grand mean.*

students in math classes (x = 59.48). However, the effect size was small indicating the magnitude of this difference is relatively small. Comparison group students' achievement did not differ based on whether they were in math or science classes (**Figure 2**). No other interaction effects were statistically significant (p's > 0.05). Boys' and girls' performance was comparable regardless of whether they were enrolled in a math or science class, and performance was also comparable across instruction groups. The three-way interaction was also not statistically significant (p > 0.05).

3.2 Predictors of math and science achievement

Two separate multiple regressions were conducted to identify significant predictors of math and science achievement. The Breush-Pegan and Koenker test was significant for both the science and mathematics scores, indicating the heteroskedasticity in the data. To correct for the non-constant variance in the regression results (heteroscedasticity), a heteroskedasticity- consistent estimator (HC3 method) was used to calculate robust standard errors in both regression analyses. The predictor variables for both analyses included three classroom factors (instruction group, disciplinary context, integration opportunities) and two student factors (ethnicity, gender). Two of the three classroom variables (instruction group and integration opportunity) were found to be significant predictors of students' science scores [$F(9, 534) = 205.378$, $p < 0.001$)] with an R^2 of 0.772 indicating the model accounted for 77.2% of the variance in science scores. Students' predicted science achievement scores was equal to 44.028–9.758 (instruction group) + 1.856 (integration opportunity) where instruction group was coded as 0 = $(MS)^2$ and 1 = control group, and integration opportunity was measured on a five-point Likert scale (1 = very few or no integration opportunities, 5 = daily integration opportunity). Results show that science scores were 9.758 points higher for students in the $(MS)^2$ group and their scores increased by 1.856 points for every one-point increase in reported integration opportunity. Neither student ethnicity nor student gender were significant predictors of the science achievement. Disciplinary context was also not a significant predictor of science achievement. Further, none of the interaction terms were significant predictors of science achievement.

Like the science achievement results, two classroom variables (instruction group, integration opportunity) significantly predicted mathematics achievement scores [$F(9, 534) = 195.794$, $p < 0.001$)] with an R^2 of 0.764 indicating the model accounted for 76.4% of the variance in student math scores. Students' predicted math achievement score was equal to 18.178 + 9.271 (instruction group) + 5.293 (integration opportunity). Instruction group and integration opportunity were coded the same as in the science achievement analyses. Math scores were 9.271 points higher for the students in the $(MS)^2$ group and increased by 5.372 points for every one-point increase in reported integration opportunities. Like science scores trends, neither student ethnicity nor student gender were significant predictors of the math achievement. Disciplinary context nor any of the interaction effects were significant predictors of math achievement.

4. Limitations

There are design limitations that should be considered when interpreting the results of this study. Since all of the $(MS)^2$ teachers showed interest in integration when enrolling in this program, there may be a selection bias for the intervention group influencing the results. Another limitation is the lack of additional information related to student background, such as prior math and science achievement. This limited our ability to test for potential initial differences between instruction groups, meaning that the generalizability for this study is limited. Initial differences between the groups cannot be fully attributed to effectiveness of the $(MS)^2$ group since the groups may not have been equivalent. Further, given the homogeneity issues noted in the MANOVA results section, the results may be biased. However, since the sample sizes of the instruction group and comparison group are comparable (265 and 279) the analysis of variance can be considered robust for violation of homogeneity of the variances [40]. Another potential study limitation is that results may not be generalizable to students outside of the region within the United States the study was conducted. Therefore, additional studies should focus on replicating this current project using a national representative sample. Despite these limitations, the large effect sizes suggest the approach is promising.

5. Discussion

This study sought to examine the extent to which integrated science and mathematics instruction impacted student science and math achievement and to explore the extent to which opportunities to engage in integrated science and mathematics instruction significantly predicted student achievement above and beyond student gender, ethnicity, and disciplinary context. The results show $(MS)^2$ positively impacted middle school students' science and mathematics knowledge in this sample of students. Students who received $(MS)^2$ outperformed students in the comparison group on science and mathematics achievement tests.

5.1 Interaction effects

The significant disciplinary context by instruction group effect for science suggests that students who received $(MS)^2$ integrated instruction in science classrooms slightly outperformed those who received $(MS)^2$ integrated instruction in math classrooms. This interaction effect was not found for mathematics achievement. This pattern suggests that the impact of integrated instruction differs for mathematics achievement depending on whether the instruction was delivered in a science or mathematics classroom. Stronger science achievement was observed in $(MS)^2$ science classrooms than $(MS)^2$ mathematics classrooms; but the mathematics achievement impact was comparable in both contexts. This finding may be a result of differences in pedagogical practices, learning objectives, or skill requirements that are different in science and mathematics classrooms. Given science classrooms would target more science content than mathematics classrooms, it is not surprising that students in science classrooms slightly outperform students who received integrated instruction in mathematics classrooms. Thus, this finding likely reflects the differences in content standards for which science and mathematics teachers are responsible. On the other hand, a disciplinary context by instruction group interaction effect for mathematics achievement was not observed. Therefore, one could argue that an interaction effect should have been observed given the emphasis in mathematics in math courses. Yet, mean scores on the mathematics ITBS were comparable whether students received $(MS)^2$ instruction in a math or science classroom.

This pattern could be attributable to the science emphasis in the $(MS)^2$ professional development experience. Teachers in the $(MS)^2$ group took more science than mathematics courses. This imbalance in content emphasis may have contributed to the lack of instructional group by disciplinary context interaction effect for mathematics achievement. Therefore, future studies should examine the impact of different proportions of science and mathematics content exposure teachers receive on their students' achievement outcomes. These differences could also be attributable to student level factors such as task orientation [41] or school level factors such as socioeconomic composition [42]. Future studies should scale up efforts to examine the role of these factors in mediating the impact of integrated instruction.

5.2 Magnitude of impact on mathematics achievement

What was not expected was the findings related to mathematics achievement. The effect size for the $(MS)^2$ impact on mathematics achievement ($\eta^2 = 0.663$) was larger than reported impacts in past research [43]. Although there is limited research on the differential impact of integrated instruction on mathematics achievement, past research has suggested that integrated instruction may be problematic for mathematics content learning [44, 45]. For example, in a meta-analysis conducted by [43], 9 of 13 studies had effect sizes for mathematics achievement of 0.2 or less. In contrast, 6 of 13 studies had effect sizes for science achievement of .50 or greater and nine studies having effect sizes above 0.20. Further, other scholars have argued that an integrated approach may not be consistent with mathematics epistemic knowledge which, in turn, restricts mathematics learning [46]. The higher impact observed for mathematics achievement may have been due to the prolonged training teachers received in the project or the multiple opportunities teachers were provided to engage with faculty in science and mathematics research laboratories. These experiences may have contributed to differences in teaching practices that contributed to greater gains in student mathematics achievement. Though, as noted above, the proportion of mathematics to science content engagement in the training may not be ideal for developing teachers' mathematics PCK.

5.3 Predictors of science and mathematics achievement

The regression results provide a potential explanation for the higher effect sizes observed in mathematics achievement. Recall that $(MS)^2$ instruction increased student mathematics and science achievement scores by 9 to 10 points. The relative consistency in the provision of integrated opportunities also significantly improved student math and science scores. We also observed the magnitude of the impact on the math scores was higher than the science scores. Integrated opportunities increased mathematics scores by over 5 points for each level of increase of integration opportunity, compared to a 1.9-point increase in science scores. Students in our sample, particularly the $(MS)^2$ group, may have had more consistent opportunities to engage in science and mathematics integrated instruction than students from other studies. Minimally, the regression results suggest that another source of variability of impact on mathematics achievement is related to the frequency of opportunities students are provided to engage in integrated lessons. It is possible the larger effect sizes are attributable to the quality of instruction students in $(MS)^2$ group received as resulting from the extensive training $(MS)^2$ teachers received. Although we would expect the intensive training teachers received would impact their commitment to provide consistent opportunities for integrated instruction, student reports of integrated instruction opportunity was a significant predictor of mathematics achievement which suggests variability in the frequency of these opportunities in the $(MS)^2$ sample. This finding underscores the impact of consistent opportunities to engage in integrated instruction is also important in raising student scores, particularly for mathematics. Future studies could examine this impact more directly by varying the consistency in the frequency of integrated opportunities.

5.4 Differences across gender and race/ethnicity

We also found that gender and race/ethnicity did not significantly predict student outcomes despite consistent research showing both gender [47] and race/ethnicity gaps [48] in math-science achievement in the US. Ethnicity and gender did not significantly predict science and mathematics achievement in the integrated instruction groups regardless of whether they were in an integrated math or integrated science class. This finding suggests that integrated science and mathematics instruction can lead to more equitable outcomes in mathematics and science achievement and therefore should be considered as a promising approach to reducing achievement gaps. These findings are encouraging given that others have found the impact is not uniform across groups of students. For instance, in a large scale, quasi-experimental study, [48] found achievement disparities persisted, advantaging Anglo-American students, despite the integrated SE instruction. [48] also found a significant relationship between quality of SE integration and student outcomes. They concluded that low integration quality largely contributed to the lack of impact on student outcomes. The authors also acknowledged that the lack of instructional scaffolds for diverse learners likely contributed to group disparities.

6. Implications

The results of this study suggest integrated science and mathematics instruction is a possible strategy for improving student math and science achievement. Given the increasing evidence supporting the argument that integrated instruction promotes critical thinking, motivation, and persistence in STEM, it is critical to support teachers and schools in developing this approach to STEM instruction. Integrated instruction appears to provide the experiences that shape student interest; over time, these opportunities can lead to sustained interests in STEM which is important in developing positive STEM identity and self-efficacy [7].

However, integrated instruction requires significant investment in the development and dissemination of quality professional development for teachers. Teachers in the $(MS)^2$ group received 2 years of professional development and they engaged with science and mathematics professors in designing curricular materials, practicing the presentation of synergistic concepts and principles, and conducting authentic disciplinary practices in university and national laboratories. This degree of investment is likely not possible in schools in the USA. Therefore, it is important for school districts to partner with university programs to provide experiences that meaningfully engage them in disciplinary practices with STEM professionals.

Further, teachers seeking to use integrated science and mathematics instruction need administrative support. In practice, this means administrators should allot time for teachers to collaborate with each other on the development of integrated lessons that work for science and mathematics classrooms. This structured time provides opportunities for science and mathematics teachers to gain from each other's expertise and experience. In addition, during appraisals and observations of teachers, administrators should also recognize teacher investment in integrated instruction and provide feedback in terms of the consistency of integration opportunities provided to students as well as the quality of math-science integration. Therefore, there is a need to expand current evaluation and observation tools to include an integrated instruction domain. For instance, the Texas Education agency (TEA) implemented the Texas Teacher Evaluation and Support System (T-TESS), which captures the quality of teacher instruction and its effectiveness in students' outcomes. This measure includes three main components: goal-setting and professional development, the evaluation cycle (informal and formal observations), and a student growth measure. Therefore, evidence of student opportunities to be engaged in integrated instruction can be collected by surveying the students' perceptions of integrated instruction including how often these opportunities were provided to them and the degree to which it helped their understanding of math and science content as a way to gauge students' perceptions of the impact of these opportunities. Past large-scale studies investigating student perceptions of teaching quality as predictors of science achievement are mixed [49, 50] however, these studies use general surveys of student perceptions. This study surveyed student perceptions of the frequency of integrated instruction, not general quality perceptions.

7. Conclusion

Overall, the results of our study suggest that integrated science and mathematics instruction can yield strong results in mathematics and science learning for all students. The large effect sizes indicate integrated science and mathematics instruction is promising in raising achievement for students from diverse ethnic and socioeconomic backgrounds, provided schools invest in supporting teachers in developing their ability to provide all students with meaningful and consistent integrated science and mathematics opportunities.

Author details

Zenaida Aguirre-Muñoz[1*], Brian Dang[2] and Elias S. Loria Garro[2]

1 Cognitive and Information Sciences, University of California, Merced, United States

2 Psychological, Health, and Learning Sciences, University of Houston, United States

*Address all correspondence to: zaguirre-munoz@ucmerced.edu

References

[1] National Council of Teachers of Mathematics (NCTM). *Principles and Standards for School Mathematics*. Reston, VA: NCTM; 2000

[2] Miranda RJ, Hermann RS. An integrated instructional approach to facilitate inquiry in the classroom. Science Scope. 2012;**35**(8):66-72

[3] NGSS Lead States. *Next Generation Science Standards: For States, by States*. Washington, DC: The National Academies Press; 2013

[4] English L. STEM education K-12: Perspectives on integration. International Journal of STEM Education. 2016;**3**(3):1-8. DOI: 101186/s40594-016-0036-1

[5] English L. Advancing elementary and middle school STEM education. International Journal of Science and Mathematics Education. 2017;**Suppl 1**: S5-S24

[6] Vasquez J. STEM: Beyond the acronym. Educational Leadership. 2015;**72**(4):10-15

[7] Aguirre-Muñoz Z, Pantoya M, Pando M, Loria EG. Engineering integration in science elementary classrooms: Effects of disciplinary language scaffolds on English learner content learning and engineering identity. Journal of Engineering Education. 2021;**110**(*3*):517-544. DOI: 10.1002/jee.20409

[8] Struyf A, De Loof H, Boeve-de Pauw J, Van Petergem P. Students' engagement in different STEM learning environments: Integrated STEM education as promising practice? International Journal of Science Education. 2019;**41**(10):1387-1407. DOI: 10.1080/09500693.2019.1607983

[9] Aguirre-Muñoz Z, Pantoya M. Engineering Literacy and Engagement in Kindergarten Classrooms. Journal of Engineering Education. 2016;**105**(4):630-654

[10] Skilling K, Bobis J, Martin AJ, Anderson J, Way J. What secondary teachers think and do about student engagement in mathematics. Mathematics Education Research Journal. 2016;**28**(4):545-566. DOI: 10.1007/s13394-016-0179-x

[11] Tang KS, Williams PJ. STEM literacy or literacies? Examining the empirical basis of these constructs. The Review of Education. 2018;**7**(3):675-697. DOI: 10.1002/rev3.362

[12] Skilling K. Student STEM beliefs and engagement in the UK: How they shift and are shaped through integrated projects. In: Anderson J, Li Y, editors. Integrated Approaches to STEM Education: An International Perspective. Switzerland: Springer; 2020. pp. *251-270*

[13] Aguirre-Muñoz Z, Yeter I, Loria E, Koca F. Building teachers' capacity to integrate math and science content: Implications for professional development and learning. Journal of Science Teacher Education. 2021;**32**(1):62-84. DOI: 10.1080/1046560X.2020.1789814

[14] Penuel WR, Reiser BJ, Novak M, McGill T, Frumin K, VanHorne K, et al. Using co-design to test and refine a model for three-dimensional science curriculum that connects to students' interests and experiences. Paper presented at the Annual Meeting of the American Educational Research Association. New York, NY; 2018

[15] Berlin DF, White AL. The Berlin-White integrated science and mathematics model. School Science and Mathematics. 1994;**94**(1):2-4. DOI: 10.1111/j.1949-8594.1994.tb12280.x

[16] Breiner JM, Harkness SS, Johnson CC, Koehler CM. What is STEM? A discussion about conceptions of STEM in education and partnerships. School Science and Mathematics. 2012;**112**(1):3-11. DOI: 10.1111/j.1949-8594.2011.00109.x

[17] Bossé MJ, Lee TD, Swinson M, Faulconer J. The NCTM process standards and the five Es of science: Connecting math and science. School Science and Mathematics. 2010;**110**(5):262-276. DOI: 10.1111/j.1949-8594.2010.00033.x

[18] Furner JM, Kumar DD. The mathematics and science integration argument: A stand for teacher education. Eurasia Journal of Mathematics, Science, & Technology Education. 2007;**3**(3):185-189. DOI: 10.12973/ejmste/75397

[19] Lederman NG, Lederman JS. Is it STEM or "S & M" that we truly love? Journal of Science Teacher Education. 2013;**24**(8):1237-1240. DOI: 10.1007/s10972-013-9370-z

[20] Huntley MA. Design and implementation of a framework for defining integrated mathematics and science education. School Science and Mathematics. 1998;**98**(6):320-327. DOI: 10.1111/j.1949-8594.1998.tb17427.x

[21] Huntley MA. Theoretical and empirical investigations of integrated mathematics and science education in the middle grades with implications for teacher education. Journal of Teacher Education. 1999;**50**(1):57-67. DOI: 10.1177/002248719905000107

[22] Czerniak CM, Weber WB, Sandmann A, Ahern J. A literature review of science and mathematics integration. School Science and Mathematics. 1999;**99**(8):210-220. DOI: 10.1111/j.1949-8594.1999.tb17504.x

[23] Hurley MM. Reviewing integrated science and mathematics: The search for evidence and definitions from new perspectives. School Science and Mathematics. 2001;**101**(5):259-270. DOI: 10.1111/j.1949-8594.2001.tb18028.x

[24] McComas WF, Wang HA. Blended science: The rewards and challenges of integrating the science disciplines for instruction. School Science and Mathematics. 1998;**98**(6):340-348. DOI: 10.1111/j.1949-8594.1998.tb17430.x

[25] Bragow D, Gragow KA, Smith E. Back to the future: Toward curriculum integration. Middle School Journal. 1995;**27**(2):39-46. DOI: 10.1080/00940771.1995.11496152

[26] Greene LC. Science-centered curriculum in elementary school. Educational Leadership. 1991;**49**(2):42-46

[27] Kim MK, Roh S, Cho MK. Creativity of gifted students in an integrated math-science instruction. Thinking Skills and Creativity. 2016;**19**:38-48

[28] Guthrie JT, Wigfield A, VonSecker C. Effects of integrated instruction on motivation and strategy use in reading. Journal of Educational Psychology. 2000;**92**(2):331-341. DOI: 10.1037/0022-0663.92.2.331

[29] McHugh L, Kelly AM, Burghardt MD. Professional development for a middle school mathematics-infused science curriculum. Journal of Science Teacher Education. 2018;**29**(8):804-828. DOI: 10.1080/1046560X.2018.1514825

[30] Orime SA, Ambusaidi A. The impact of using the integration approach between science and math on acquiring the skills for solving scientific problems for fourth grade students. Journal of Turkish Science Education. 2011;**8**(2):9-22

[31] Shulman LS. Those who understand: Knowledge growth in teaching. Educational Researcher. 1986;**15**(2):4-14. DOI: 10.3102/0013189X015002004

[32] Silverstein SC, Dubner J, Miller J, Glied S, Loike JD. Teachers' participation in research programs improves their students' achievement in science. Science. 2009;**326**(5951):440-442. DOI: 10.1126/science.1177344

[33] Dunbar S, Welch C. *ITBS* Research and Development Guide. Orlando, FL: Houghton Mifflin in Harcourt Publishing Company; 2015

[34] Lee G, Dunbar SB, Frisbie DA. The relative appropriateness of eight measurement models for analyzing scores from tests composed of testlets. Educational and Psychological Measurement. 2001;**61**(6):958-975. DOI: 10.1177/00131640121971590

[35] Loyd BH, Forsyth RA, Hoover HD. Relationship of elementary and secondary school achievement test scores to later academic success. Educational and Psychological Measurement. 1980;**40**(4):1117-1124

[36] Ansley TN, Forsyth RA. Relationship of elementary and secondary school achievement test scores to college performance. Educational and Psychological Measurement. 1983;**43**(4):1103-1112. DOI: 10.1177/001316448304300419

[37] Qualls AL, Ansley TN. The predictive relationship of achievement test scores to academic success. Educational and Psychological Measurement. 1995;**55**:485-498

[38] Howell DC. Statistical Methods for Psychology. Belmont, CA: Thompson Wadsworth; 2007

[39] Cohen J. Statistical Power Analysis for the Behavioral Sciences. New York: Academic Press; 1969

[40] Ramashandran KM, Tsokos CP. Mathematical Statistics with Applications in R. 2nd ed. San Francisco: Academic Press; 2015. DOI: 10.1111/j.1949-8594.2007.tb17769.x

[41] Yerdelen S, Sungur S. Multilevel investigation of students' self-regulation processes in learning science: Classroom learning environment and teacher effectiveness. International Journal of Science and Mathematics Education. 2019;**17**(1):89-110

[42] McConney A, Perry LB. Science and mathematics achievement in Australia: The role of school socioeconomic composition in educational equity and effectiveness. International Journal of Science and Mathematics Education. 2010;**8**(3):429-452

[43] Becker K, Park K. Effects of integrative approaches among science, technology, engineering, and mathematics (STEM) subjects on students' learning: A preliminary meta-analysis. Journal of STEM Education: Innovations and Research. 2011;**12**(5):23-37

[44] Baldinger EE, Staats S, Covington Clarkson LM, Contreras Gullickson E, Norman F, Akoto B. A review of concptions of secondary mathematics in integrated STEM education: Returning voice to the silent M. In: Anderson J, Li Y, editors.

Integrated Approaches to STEM Education: An International Perspective. Switzerland: Springer; 2020. pp. 67-90

[45] Tytler R. STEM education for the twenty-first century. In: Anderson J, Li Y, editors. Integrated Approaches to STEM Education: An International Perspective. Switzerland: Springer; 2020. pp. 45-66

[46] Honey M, Pearson G, Schweingruber H. *STEM* Integration in K-12 Education: Status, Prospects, and an Agenda for Research. Washington, DC: National Academies Press; 2014

[47] Stout JG, Dasgupta N, Hunsinger M, McManus MA. STEMing the tide: Using ingroup experts to inoculate women's self concept in science, technology, engineering, and mathematics (STEM). Journal of Personality and Social Psychology. 2011;**100**(2):255-270. DOI: 10.1037/a0021385

[48] Guzey SS, Harwell M, Moreno M, Peralta Y, Moore TJ. The impact of design-based STEM integration curricula on student achievement in engineering, science, and mathematics. Journal of Science Education and Technology. 2017;**26**(2):207-222. DOI: 10.1007/s10956-016-9673-x

[49] Zhang F, Bae L, Broda M. Science self-concept, relatedness, and teaching quality: A multilevel approach to examining factors that predict science achievement. International Journal of Science and Mathematics Education. 2021;**20**:503-529. DOI: 10.1007/s10763-021-10165-2

[50] Zhang D, Campbell T. An examination of the impact of teacher quality and "opportunity gap" on student science achievement in China. International Journal of Science and Mathematics Education. 2015;**13**(3):489-513

Chapter 5

Grade 10 Girls' Experiences in Choosing STEM Subjects in Rakwadu Circuit, South Africa

Israel Kibirige and Shapule Edith Modjadji

Abstract

The lopsided participation of females in science, technology, engineering, and mathematics (STEM) professions is an issue of global concern. Very few girls choose to study sciences in secondary schools in South Africa. Understanding girls' experiences in choosing science subjects may assist various education stakeholders to ensure that their roles motivate more girls to choose sciences. This study explored grade 10 girls' experiences in choosing STEM subjects. A case study was designed using 10 girls out of 145 who had chosen to study STEM subjects from three secondary schools in Limpopo Province, South Africa. Data were collected through semi-structured interviews and were analyzed thematically. Five themes emerged regarding Grade 10 girls' experiences in choosing to study STEM subjects—self-determination, anticipated value, the class environment, home influence, and social influence. Parental guidance of "girl-child" was very limited. The findings highlight that many girls in rural schools in Limpopo did not choose STEM subjects in Grade 10. These findings have far-reaching implications for all education stakeholders in the country and beyond.

Keywords: motivation, support, gender, self-determination, performance

1. Introduction

For years, countries have been concerned with the number of female students studying science, technology, engineering, and mathematics (STEM) subjects in secondary schools [1]. To study STEM, learners need to study physical science, which includes physics and chemistry. Learners who study physical science are encouraged to take mathematics and technology at high school as part of STEM subjects. There are worldwide initiatives to enhance learners' interest in STEM subjects [2], yet few girls choose to study sciences [3, 4]. Even though boys and girls have equal opportunities to study physical science, there are gender differences that influence subject choices in secondary schools and ultimately STEM careers. Although motivation in schools is important, it is often overlooked [5], and the factors that motivate girls to study physical science are not well studied and remain an area of concern [6, 7].

In South Africa, all subjects in Grades 7 to 9 are compulsory, including natural science, which incorporates physical science, life sciences, and earth sciences [8].

In Grade 10, learners choose subjects they wish to pursue up to Grade 12. It is a stage that defines the path toward STEM careers they want to pursue [4]. King and Glackin in their study [9] have shown that most students develop interest and attitudes toward STEM subjects at the age of 14. As a result, exposure to STEM subjects at this age may be crucial in shaping attitudes and interests. Researchers during teaching practice sessions noticed a very small number of girls studying physical science, which henceforth is referred to as STEM subjects, for Grade 10. The first author, a STEM teacher and a lady, was concerned with the few girls to study physical science. Girls do not choose physical science, and this alienates them from STEM careers. Once they decide not to choose physical science in secondary school, it may be difficult for them to enter a STEM degree in tertiary institutions [10]. It is no wonder the low participation of girls and women in STEM is a never-ending story [6].

Although total enrolment of girls in schools has increased [11], fewer girls than boys choose physical science in South African secondary schools [12]. This low enrolment in physical science can be partly explained by the girls' poor performance in sciences. The trends in mathematics and science study (TIMMS) [13] show that girls' science performance was poor [13], and this situation has not improved. Bottia et al. [14] attributed the poor performance to girls' attitudes, interests and while Tzu-Ling [15] attributed it to motivation toward STEM subjects. Studies suggest a lack of role models [16, 17], lack of information about STEM [18], females' lack of confidence in sciences [19], and the lifestyles related to gender [20]. Also, a few women scientists can encourage girls to study STEM subjects [21]. Finally, the low numbers of girls studying STEM subjects ultimately result in few females in STEM careers [22, 23]. One wonders what could be the challenges. How can those challenges be overcome? It was envisaged that understanding girls' experiences in STEM subjects in secondary schools could shed light on the surrounding challenges for stakeholders to identify possible solutions [1]. In South Africa, culture and the environment influence girls' choices of subjects to study. Secondary school learners in their teens show gender differences in their behaviors [23, 24] and choices. It is most likely that these differences in masculinity and femininity manifest where more boys than girls choose STEM subjects, thus sustaining the hegemony of male stereotypes [25–27]. Sekuła et al. [28] contend that females in STEM are like strangers or intruders of the male-dominated terrain. While numerous studies have identified factors that affect girls' decisions to pursue STEM subjects [15, 29–31], the findings have not been exhaustive, and some factors may be context-specific. Girls' experiences in choosing and learning STEM subjects in rural areas of South Africa are unknown. There is no published work on South African Grade 10 girls' experiences regarding choosing STEM subjects. The study explored Grade 10 girls' experiences of choosing physical science (a STEM subject) to narrow this gap. To achieve the above purpose, the research posed the following question—What are the experiences of girls in studying STEM in rural secondary schools of Limpopo, South Africa? Also, there were probe questions—What attracted you to choose science? What help did you get from your parents? Do you have a STEM female role model in your school or community? What career do you like to take? What challenges do you experience when studying science?.

2. Literature review and theoretical framework

STEM subjects are fundamental for developing national economies, yet the performance in mathematics and sciences that lead to STEM has been poor for the South

African learners [32]. The situation is worse for the secondary school girls who perform poorly in STEM and do not choose the subject. In addition, girls who perform better in science do not choose STEM subjects, hence causing a leakage of girls leaving STEM [33].

2.1 Gender disparities in STEM

Gender differences continue to exist in participation in STEM subjects (Catalyst, 2019), where many girls do not choose STEM subjects due to negative attitudes toward the subjects [23, 25]. Comparatively, boys show more positive attitudes toward science than girls [15]. The gender disparity contrasts the United Nations Sustainable Development Goal (SDG-4), which requires that all boys and girls be at the same level in accessing quality primary and secondary education by 2030 [34]. Judging from the current state of affairs in education, this may not be achieved. Furthermore, the SDG-5 necessitates gender equality to empower all girls and women in the education sector. As suggested by Kind et al. [35], gender disparity in STEM can be attributed to attitudes toward science demonstrated through seven tenets—(a) learning science in school, (b) practical work in science, (c) science outside of school, (d) importance of science, (e) self-concept in science, (f) future participation in science, and (g) combined interest in science. A study conducted in South Africa found that boys were more interested in studying STEM subjects than girls [36]. In addition, the choosing of STEM subjects may be attributed to cultural and social factors, school science curriculum, or people's perceptions toward STEM subjects [37]. The gender disparity in STEM is a multi-faceted issue that needs all stakeholders to work together to change the gender gap in STEM subjects at the secondary school level and indeed at all other levels of education. It is no wonder it involves two worlds. First, the private and the public. The private comprise families and the educational institutions that enhance skills and knowledge. Subtly, it is a place where perceptions regarding traditional gender roles are strengthened. Second, the public domain comprises the workplace, which unfortunately encourages male–female gender roles [38].

2.2 Learners' performance in STEM

Learners' poor performance is a persistent challenge in Limpopo, South Africa [24, 39]. Although the number of girls in physical science has increased in recent years [11], the number of girls choosing physical science in South African secondary schools is far less than the number of boys [12], and indeed both in developed and less developed countries [40]. The low percentage suggests that most girls are not motivated enough to study science subjects, resulting in poor performances [41–44]. Conversely, Stoet and Geary [45] show that boys and girls perform equally well in STEM subjects. Notwithstanding motivation and good grades, girls may not choose STEM subjects due to personal (micro-level), family and societal (mezo-level), and cosmopolitan culture (macro-level) reasons [45]. These three cover all spheres of a learner and spill in the careers aspirations. For example, apart from personal issues, family and institutional differences exist. Some families are more inclined to study STEM subjects than others [46], although this may vary from context to contest [47]. Studies in the United States of America (USA) support that family differences exist. For example, if a girl is first or last born in the family has different results as far as STEM subjects are concerned. The treatments children receive in the family have a bearing on their performance in school subjects. Hence, the position of the siblings and parental preferential treatments have an impact on STEM performances.

2.3 Learner enrolment

The unequal participation of girls in STEM subjects has remained a global challenge. In France, girls constitute 44.2% of physical science learners [46]. In the United States of America, the Girls, Mathematics and Science Partnership (GMSP) handled matters dealing with girls' participation in science [48]. Similarly, in Malawi, Ghana, Nigeria, and South Africa, there are gender disparities regarding learners' participation in scientific and technological subjects [49]. In Africa, 22% of girls attend secondary school and only 10% of the 22% study science [50]. This implies few girls study sciences and few could enroll in universities and take careers in STEM [22].

2.4 Factors influencing girls' choices of science subjects

Considerable literature has been published on factors influencing girls not to choose science subjects. The factors include lack of role models [51–54], lack of information about sciences, and scientific careers for learners in rural areas [55, 56]. Girls' lack of personal efficacy in science careers [57–59] attests that female role models can inspire girls to develop an interest in science careers.

Although countries differ in their social and economic status, they all experience gender differences. These differences are stratified in all levels of growth and development. In this study, the researchers focused on secondary schools. They are adolescents who are soon to leave childhood and join adulthood. Learners at this level are at crossroads. They require guidance in the now and the future choices. The researchers are reminded of the type of education that is offered. It is narrow and does not cater to the present and the future. It does not deal with the whole body, mind, dimensions and spiritual [60]. It implies that the narrowness of mind may influence girls' choices in STEM. Other factors include gender stereotypes content and teaching styles that elevate males over females [27, 61, 62]; differences in aspiration where many boys aspire and choose STEM subjects because few girls choose STEM subjects; teaching methods that favor boys and not girls [63, 64]; individual beliefs and family friends [45, 65]; school subject environment [27], and future career aspirations [15]. In summary, these factors are on three levels—a personal (micro), a family, school and friends (mezo), and cosmopolitan or ambient culture (macro).

2.5 Theoretical framework and learner experiences in STEM

Two theories guided the study—1) the Social Cognitive Theory (SCT), [66], and 2) the Situated Expectancy-Value Theory (SEVT) [67]. SCT describes self-efficacy, outcome expectancies, and goals constructs [67, 68]. It is a triadic model comprising three tenets—reciprocal causation, individuals as actors, and environmental products. Thus, SCT describes behavioral changes that an individual makes. The girls' experiences reflect a behavioral change to study STEM in this study.

The Situated Expectancy Value Theory (SEVT) [69] extends the work of Eccles [70] in dealing with choice making. SEVT has five key elements, which are as follows:

1. Individuals are motivated by achievement-related choices,

2. Proximal social cognitive aspects and dealing with within and between individual decision making is based on experiences.

3. Individuals' experiences and interpretation of experiences guide their choices,

4. Social and experiential, the cognitive, affective, and behavioral components influence individuals choice,

5. Choices are limited by prior experiences, cultural values, norms, and individuals' characteristics. In choice-making, SEVT is robust because it is situation-specific and based on cultural norms.

Girls' choices to study STEM subjects in South Africa are guided by various factors, including the situation and the culture, to relate their experiences regarding STEM. Thus, these two theories were selected because they deal with the individual's situated environment that guides behavioral changes. In their teens, high school learners are showing gender differences in their behaviors [23, 24]. It is most likely that these differences in masculinity and femininity manifest in the subject choices where more boys than girls choose physical science, thus, sustaining the hegemony of male stereotypes. Girls' experiences in choosing and learning STEM subjects in rural areas of South Africa are scanty. Therefore, this study contributes to understanding girls' experience in choosing STEM, which could interest politicians, researchers, academics, and education stakeholders to ameliorate the situation.

3. Methodology

3.1 Design

This study utilized an exploratory case study design to investigate Grade 10 girls' experiences in choosing to study STEM subjects. According to Cohen et al., [71], a case study is beneficial because it draws data from people's experiences and practices. A purposive sample [72] of 10 Grade 10 girls (age 14–16) from three schools in Rakwadu Circuit, South Africa, was used based on their choices to study STEM subjects.

3.2 Sample

Grade 10 girls from three schools, A, B, and C (4, 3, and 3) were selected. Learners one to four from school A were coded as L1A to L4A, learners one to three from school B were coded as L1B to L3B, and learners one to three from school C were coded as L1C to L3C. The three schools had 216 learners in grade 10, 145 were girls, and only 10 chose to study physical science. In this case, only 10 girls chose to study STEM subjects leaving out most of them (135) to study other subjects. For ethical considerations, all minor participants were issued with consent letters to be signed by their parents/guardians to allow their children to take part in the study. Permissions were granted from schools, the Circuit Education office, and the University of Limpopo Research Ethics Committee.

3.3 Data collection

Data were collected through semi-structured interviews. Semi-structured interviews [72] were used because they offered the interviewer a chance for in-depth discussions, follow-ups, and probing questions to clarify the responses [73]. All interviews were audio-taped, and each interview lasted for one hour, which was enough without causing fatigue to the learners [71]. Harm was avoided by explaining that the study had no impact on their academic performance and that learners could at any time withdraw from the interviews [74, 75]. Member check was performed with the participants to ensure that the captured information correctly reflected their views [76].

3.4 Data analysis

Data from the interviews were analyzed thematically to provide descriptions of the findings [44, 77]. The thematic analysis process involved identifying patterns across data sets that were important in describing a phenomenon associated with the research questions [78]. The interviews were transcribed verbatim, and the transcripts were read line by line several times to gain insights into the participants' responses. The researchers generated a codebook to make themes based on the theories and collected data [79, 80]. In theory-based, two researchers and one expert coded the data and compared codes. All three researchers used similar codes to form categories, and the last categories were organized into themes [81]. Where there were disagreements, a consensus was reached using the inter-observation agreement [82] formula, where agreements were divided by the sum of agreements and disagreements. The product was multiplied by 100%, and a value of 90% was appropriate for this study. Thus, a codebook was used to analyze data deductively, while the collected data were analyzed inductively, where the researchers read paragraph by paragraph to find out the general pattern.

4. Results

The girls' responses are categorized into five major themes—personal factors, anticipated value, class environment, home influence, and social influence. The themes are presented below with exemplars of comments from the participants.

4.1 Theme 1: Self-determination

Self-determination included positive attitudes, interest in the subject, and performing well. When learners were asked why they chose STEM subjects, they indicated that physical science was an interesting subject they enjoyed. Two sample excerpts from participants:

L1A: "*Physical science is interesting, and I enjoy it. I understand science concepts.*

L2B: "I always wanted to study physical science. People say it is a difficult subject, but I find it to be easy. Unlike mathematics, physics is simple, and I understand it better than mathematics. I perform well in the tests and assignments. Physics is an enjoyable subject."

The girls expressed determination to take on science careers, where physical science was a prerequisite. Participant L4A explained: *"I chose physical science because I want to be a Medical Doctor. I must have physical science as a subject because it is a prerequisite for entrance into Medicine."*

One participant indicated that whereas her father wanted her to be a nurse, she was determined to study hard to become an electrical engineer.

L3C: "I chose physical science because I want to do civil engineering at the university."

4.2 Theme 2: Anticipated value

All the 10 study participants indicated that they were motivated by future careers to study physical science. All participants stated that physical science was imperative for STEM careers (**Table 1**).

L3C: I know there are opportunities for well-paid jobs when one does sciences. I can secure a scholarship for further studies".

I think I will get a good-paying job. L2C: "I think our lives would not be the same if people were not studying physical science because people who invent things are scientists."

4.3 Theme 3: The class environment

The majority of participants indicated that they received continuous support from educators. L3B stated: "*My teachers encouraged me to choose physical sciences and mathematics since I performed well".*

L4A*: "Our teacher is friendly and wants us to succeed. He provides extra time to complete our work.*

All the participants appreciated the role of group work in learning physical science.

L3A: "*Working as a team helps us to grasp concepts.*"

They also singled out some discouraging classroom experiences.

L2A stated: "It is discouraging when the teacher concentrates on those who understand concepts faster than others, those who are smarter." This was further highlighted by L3C: "*If teachers consider you to be a slow learner or less intelligent, they do not give you much time, and sometimes they can insult you with words like...maths and physical science are for smart students... Such words discourage, but because I love Physics, I will work hard.*"

Other disobliging experiences included a lack of resources, such as laboratories, science equipment, computer centers, and an internet connection, which made learning physical science hard. *L1A said: It is difficult to learn physical science in classes where there is no science equipment. There is no laboratory to do practical work at our school, and it is sometimes difficult to understand concepts. However, I continue learning science because I enjoy it.*

L1C: "But the challenge in my school is learning science without doing experiments."

All participants from the three schools lacked laboratories, libraries, or had no access to the internet.

Future career	**Number**	**(%)**
Engineer	2	20
Doctor	2	20
Pilot	2	20
Pharmacist	1	10
Optometrist	1	10
Biotechnologist	1	10
Nurse	1	10
Total	10	100

Table 1.
Careers for learners.

4.4 Theme 4: Home influence

The majority (eight out of the 10) study participants showed that they did not get help from the family when choosing subjects to study or doing physical science assignments at home. Of the 10 participants, only two (20%) received some help from family members (**Table 2**).

Table 2 indicates that only 20% of parents/family members played a role in the girls' choosing of physical science. Excerpts from participants:

> *"My father is a Teacher, when I do not understand some of the things or questions he helps me. But he does not know most of the things because he is not a physical science teacher... My parents motivated me to choose the science stream. My father wanted me to do actuarial science but I want to be doctor."*

One participant's father wanted her to become a nurse.

The two participants who declared to have received family support had some educated members at home; other girls indicated their parents did not have much education.

L1A: "*My parents do not know science. My mother did not study science and my elder brother completed Grade seven.*"

L2B: "My parents passed away and I had no one to assist me because I am the eldest in the family and I have to take care of my siblings."

L3C (whose parents were migrant workers): "*No one helps me with my school work. When I come home, I have to fetch water, clean the house, and cook. When I finish my chores, I study and write homework.*"

L2A: "*When I come back home from school, I have to do house chores.*"

4.5 Theme 5: Social influence

Learners indicated the influence of role models within the community was important. Teachers of STEM subjects can also be role models for high school learners to emulate. Few role models, such as a medical doctor, friends, and teachers, were reported here below:

L1C "I want to become a medical doctor because there were role model doctors in the community. They are my role models to emulate."

L3A: "My friends who are not doing physical sciences say it is a difficult subject. Those who are in my science class, we help each other every time we have tasks to do at home. Sometimes we do our homework together here at school."

L3C: "*I chose physical sciences because I want to do civil engineering at the university, I attended career guidance and it was interesting to see what civil engineers do.*"

L4A*: "Our teacher is a nice person. He always wants the best from us. When we do not understand something he stays with us so that we can understand."*

Grade	Support	%	No support	%
Grade 10	2	20	8	80

Table 2.
Support received from learners' homes.

5. Discussion

The study explored Grade 10 girls' experiences in choosing physical science in South Africa. The study established that girls who studied physical science in Grade 10 were very low in the selected schools. Five themes from girls' experiences to choose STEM subjects were self-determination, anticipated value, class environment, home environment, and social influence. The study participants expressed a positive attitude and interest in science. The positive attitudes of girls in physical science contradict studies that allude to girls' negative attitudes toward science [83].

The girls' choices of physical science indicated self-determination. It is no wonder they exhibited positive attitudes toward the subject. Machingambi [84] suggests that positive attitudes may affect performance, while negative attitudes may lead to a lack of interest. The girls' excellent performance increased their confidence to choose physical science, suggesting that girls in South Africa are guided by the situation and the culture to choose STEM subjects. These observations agree with the Situated Expectancy-Value Theory (SEVT), where self-determination abetted girls' interest in STEM subjects to break the social norm of not choosing STEM [69].

The study findings are consistent with DeWitt [85], who concluded that girls who held science aspirations perform well. Archer et al. [86] concluded that "science capital," which includes economic, social, and cultural capital that relates to science would be necessary to fill the gap of the less represented females in STEM. Thus, learners may have to develop inner confidence, positive beliefs, and environmental contexts regarding their academic abilities [87–89]. These findings also align with the Social Cognitive Theory (SCT) concerning the learners' environment, where it is postulated that science, in most cases, is for males. The observed mismatch between femininity and science is a well-known fact that negatively impacts girls [86]. Girls at 13 change their attitudes toward science, exacerbating gender parity [72]. Despite popular gender stereotyping, the girls in the study expressed self-determination in pursuing physical science to get into predominantly masculine STEM professions [90]. Thus, the social aspects are clear in the three tenets of SCT—1) the personal, which operates at an individual level; 2) the socialization of an individual within the environment; and 3) the collective level, where all people work in unison to shape the decisions in their societies [91]. All these three tenets apply to learners who are social beings that make choices regarding the subjects to study in high school.

All the girls in the study had chosen STEM subjects, and their choices were implied in anticipation of lucrative jobs if they pursued STEM careers. The findings correlate with Mghweno et al. [92], who contend that career is a determinant factor in high school subject selection. However, the finding of girls' deliberate choices contradicts Dabula and Makura [93], who showed that career choices for many secondary school learners were accidental and were imposed by external forces in the South African context.

While the study participants pointed out some aspects in the classroom that motivated them to choose physical science, such as support from teachers and peers, many negative experiences were dissuading. Some of the negative influences included educators' scornful remarks and the lack of vital science resources. All the schools that participated in the study did not have laboratories and lacked basic science equipment, libraries, and internet connectivity. These poor resources disadvantaged learners because they did not develop practical skills. Despite the lack of resources, Kibirige and Bodirwa [94] show that scientific investigations can be done using technology to increase learners' interests and learning outcomes. With the increase in technology, it may be possible for girls to cope with science without proper physical resources. Our observations agree with the Social Cognitive Theory (SCT), which deals with an individual and the environment. The effect of the school-based factors agrees with Anders et al. [95], who found that in England, the type of school environment learners finds themselves in played a significant role in choosing subjects. Thus, school environment factors, such as curriculum, teachers, level of resources, and structures, may motivate or demotivate girls from choosing STEM subjects in high school [96].

Besides school factors, the home environment affected some girls' choices of physical science. In this study, only 20% of the parents supported girls in choosing sciences and could assist them with homework. The low family support can be attributed to the social and economic characteristics of the parents. Although research in the United States indicated that socio-cultural factors influence girls' participation in science [97], Ramnarain [98] in South Africa views personal (intrinsic) and external (extrinsic) factors that are associated with the Social Cognitive Theory (SCT) as integral parts of science inquiry learning.

Furthermore, Mujtaba and Reiss [96] asserted that significant factors are associated with extrinsic motivation. For instance, some girls indicated that they chose STEM subjects because they wanted to be like female doctors who were their role models in the community. Considering their reasons for choosing STEM subjects, girls in the study perceived the critical value of science, which may have motivated them. This finding is consistent with Hyde and Janet [97] and Wise and Simmons [99], who indicated that learners acknowledged the value of science. Thus, the quantity and quality of the content may enhance learners' interest and increase their self-efficacy [8]. Research from Greece shows that teachers can exert influence on learners to gain interest in STEM subjects. Studies show that pre- and primary school learners can be taught STEM subjects because they can comprehend science concepts more than anticipated [100–102]. Early learners' exposure to STEM increases their chances of espousing STEM careers [100]. Chatzopoulos et al. [103] contend that using DuBot based on Action Research, using visuals on a tablet, smartphone, and personal computers, and using low-cost materials can motivate learners. These types of innovations are useful for STEM teachers to emulate to enhance motivation of their learners to choose STEM careers and contribute to narrowing the gap between genders [101]. Unfortunately, despite the positive intentions of the teacher to use STEM methods, there are few teachers in pre- and primary schools and high schools who use STEM methods to teach science [104].

As the gender gap persists in STEM subjects, Marie et al. [105] contend that the focus should be on identifying factors that influence the girls' career choices and developing relevant programs that enhance girls' interest in STEM subjects. Career preparation in secondary schools is essential for career development [106] because learners align their subjects with the anticipated career [4]. Interventions should focus on lower grades to avoid girls' leakage at Grade 9 in South Africa. Notwithstanding the huge numbers of girls in Grade 10 that did not choose STEM subjects in Limpopo, there is a need to find out if this scenario reflects a national trend. Thus, more studies are needed to identify why many girls do not choose STEM subjects. The findings of this study have far-reaching implications for all educational stakeholders, such as subject teachers, curriculum advisers, textbook authors, to include relevant materials for the "girl-child" to be motivated to choose STEM subjects.

6. Limitations

The limitation was the small sample of Grade 10 girls from a rural area in South Africa. Therefore, the findings cannot be generalized. The study could be replicated using qualitative and quantitative approaches with larger samples of girls in rural, semi-urban, and urban areas. Girls who did not choose STEM subjects and teachers who were not interviewed in this study could be included in future studies to corroborate learners' responses. Despite those limitations, the findings from this study render credence to girls' experiences in choosing to study STEM subjects in South Africa.

7. Conclusion and recommendations

The study reveals that the experiences and factors that motivate girls to choose to study STEM subjects are diverse. They included self-determination, aspirations, anticipated value, the class environment, home environment, and social influence. According to Almukhambetova and Kuzhabekova [45], these factors can be summed into three general levels—micro, mezo, and macro. How can we improve girls' choices to study STEM subjects? How can we assist girls to improve their aspirations? Since the gender gap or disparity in STEM is a global challenge, which method can be applied that will suit all nations? These questions provoke humanity to look for real-life solutions. A one-man and a single approach may be futile. Therefore, a team of education stakeholders equipped with multi-faceted approaches is necessary. These approaches will have significant implications for STEM teachers in the country and beyond.

The study recommends that the interventions must be done at the school level to support learners in lower grades with career guidance, for science teachers to affirm learners' self-efficacy, and for policymakers to guarantee the availability of the science resources that make science learning more interesting. Educators need to be equipped with skills to support learners emotionally and academically to make STEM subjects attractive. Also, parents need to be sensitized to increase their involvement in "girl-child" education.

Acknowledgements

The authors thank teachers of various schools who assisted in arranging for time to interact with the learners.

Notes/thanks/other declarations

Thanks to the participating schools.

Author details

Israel Kibirige* and Shapule Edith Modjadji
University of Limpopo, Sovenga, South Africa

*Address all correspondence to: israel.kibirige@gmail.ac.za; kibirigei@gmail.com

References

[1] Cheryan S, Ziegler SA, Montoya AK, Jiang L. Why are some STEM fields more gender balanced than others? Psychological Bulletin. 2017;**143**(1):1-35. DOI: 10.1037/bul0000052

[2] Hattie J. Visible Learning: A Synthesis of over 800 Meta-Analyses Relating to Achievement. London, England: Routledge; 2009

[3] Lazowski RA, Hulleman CS. Motivation interventions in education: A meta-analytic review. Review of Educational Research. 2016;**86**(2):602-640. DOI: 10.3102/0034654315617832

[4] Ardura D, Pérez-Bitrián A. The effect of motivation on the choice of chemistry in secondary schools: Adaptation and validation of the science motivation questionnaire II to Spanish students. Chemistry Education Research and Practice. 2018;**19**(3):905-918. DOI: 10.1039/C8RP00098K

[5] Department of Basic Education (DBE). National Diagnostic Report on Learner Performance, 2014. South Africa: National Department of Basic Education; 2014. Available from: http:// www. education.gov.za

[6] Koch J, Irby B, Polnick B, editors. Girls and Women in STEM: A Never Ending Story. IAP; 2014

[7] Usher A, N, Kober. Student motivation – An overlooked piece of school reform. In: Centre on Education Policy, Graduate School of Education and Human Development. The George Washington University; 2012. Available from: https:// files.eric.ed.gov/fulltext/ED532666.pdf

[8] Du Plessis E. Introduction to CAPS: Curriculum and Assessment Policy Statement. Pretoria: UNISA; 2013

[9] King H, Glackin M. Supporting science learning in out-of-school contexts. In: Osborne J, Dillon J, editors. Good Practice in Science Teaching: What Research Has to Say. New York: McGraw-Hill Education; 2010. pp. 259-270

[10] van Langen AV, Dekkers H. Cross-national differences in participating in tertiary science, technology, engineering and mathematics education. Comparative Education. 2005;**41**(3):329-350. DOI: 10.1080/03050060500211708

[11] Webb P. Towards an integrated learning strategies approach to promoting scientific literacy in the south African context. International Journal of Environmental & Science Education. 2009;**4**(3):313-334. Available from: https:// files.eric.ed.gov/fulltext/EJ884400.pdf

[12] Ngila D, Boshoff N, Henry F, Roseanne D, Malcom S, Thomson J. Women's representation in national science academies: An unsettling narrative. South African Journal of Science. 2017;**113**(7/8) Art. 2017-0050, 7 pages:2012. DOI: 10.17159/sajs.2017/20170050

[13] Reddy V, Visser M, Winnaar L, Arends F, Juan AL, Prinsloo C, et al. TIMSS 2015: Highlights of Mathematics and Science Achievement of Grade 9 South African Learners. Commissioned by the Department of Basic Education; 2016. Available from: http://hdl.handle. net/20.500.11910/10673

[14] Bottia MC, Stearns E, Mickelson RA, Moller S. Boosting the numbers of STEM majors? The role of high schools with a STEM program. Science Education. 2018;**102**(1):85-107. DOI: 10.1002/ sce.21318

[15] Tzu-Ling H. Gender differences in high-school learning experiences,

motivation, self-efficacy, and career aspirations among Taiwanese STEM college students. International Journal of Science Education. 2019;**41**(13):1870-1884. DOI: 10.1080/09500693.2019.1645963

[16] Sikhwari TD. Investigating first year students' abilities to make appropriate career choices. International Journal of Educational Science. 2015;**8**(2):363-368 10.1080/ 09751122.2015.11890258

[17] Trotman A. Why don't European girls like science or technology? 2017. Retrieved November 12, 2019 from https://news.microsoft.com/europe/features/dont-european-girls-like-sciencetechnology/ UNESCO

[18] Mwantimwa K. What motivates students' decisions on programmes to pursue at university level: The role of information and knowledge. Higher Education. 2021:1-19. DOI: 10.1007/ s10734-021-00698-4

[19] Ndebele C. Gender and school leadership: Breaking the glass ceiling in South Africa. Multidisciplinary Journal of Gender Studies. 2018;**7**(2):1582-1605. DOI: 10.17583 /generos.2018.2438

[20] Eccles JS, Wang MT. What motivates females and males to pursue careers in mathematics and science? International Journal of Behavioral Development. 2016;**40**(2):100-106. DOI: 10.1177/0165025415616201

[21] Canning EA, Harackiewicz JM. Teach it, don't preach it: The differential effects of directly communicated and self-generated utility-value information. Motivation Science. 2015;**1**:47-71. DOI: 10.1037/mot0000015

[22] Shumba A, Naong M. Factors influencing students' career choice and aspirations in South Africa. Journal of Social Sciences. 2012;**33**(2):169-178. DOI: 10.1080/09718923.2012.11893096

[23] Naiker M, Sharma B, Wakeling L, Johnson JB, Mani J, Kumar B, et al. Attitudes towards science among senior secondary students in Fiji. Waikato Journal of Education. 2020;**25**(1):57-72. DOI: 10.15663/ wje.v25i0.704

[24] van der Berg S, van Wyk C, Burger R, Kotze J, Piek M, Rich K. The Performance of Low Fee Independent Schools in South Africa - What Can Available Data Tell? Stellenbosch: Stellenbosch Economic Working Papers, 1; 2017. DOI: 10.2139/ ssrn.2973229

[25] Ertl B, Luttenberger S, Paechter M. The impact of gender stereotypes on the self-concept of female students in stem subjects with an underrepresentation of females. Frontiers in Psychology. 2017;**8**:703. DOI: 10.3389/fpsyg.2017.00703

[26] Neuhaus J, Borowski A. Self-to-prototype similarity as a mediator between gender and students' interest in learning to code. International Journal of Gender, Science and Technology. 2018;**10**:233-252

[27] Makarova E, Aeschlimann B, Herzog W. The gender gap in STEM fields: The impact of the gender stereotype of math and science on secondary students' career aspirations. Frontiers in Education. 2019;**4**:60. DOI: 10.3389/feduc.2019.00060

[28] Sekuła, P., Struzik, J., Krzaklewska, E., & Ciaputa, E. Gender dimensions of physics. A Qualitative Study from the European Research Area. GENERA n. 665637. Jagiellonian University. 2018

[29] Clegg A. Girls and Science: A Training Module on Motivating Girls to Embark on Science and Technology

Careers. 2010. Available from: https://genderit.org/resources/girls-and-science-training-module-motivating-girls-embark-science-and-technology-careers

[30] Nyarko K. The influence of peer and parent relationships on adolescents' self-esteem. IFE PsychologIA: An International Journal. 2012;**20**(2):161-167. Available from: https://www.ingentaconnect.com/content /sabinet/ifepsyc/2012/00000020/00000002/art00020

[31] Thornton A. Gender equality inSTEMis possible. These countries prove it. World Economic Forum. Available from: https://www.weforum.org/agenda/2019/03/gender-equality-in-stem-is-possible/

[32] Letsoalo ME, Masha JK, Maoto RS. The overall performance of grade 12 mathematics and physical science learners in South Africa's Gauteng Province. African Journal of Gender, Society & Development. 2019;**8**(1):9-42

[33] van den Hurk A, Meelissen M, Langen V. Interventions in education to prevent STEM pipeline leakage. International Journal of Science Education. 2019;**41**(2):150-164. DOI: 10.1080/09500693.2018.1540897

[34] AmirthamA S, Kumar A. Gender parity in STEM higher education in India: A trend analysis. International Journal of Science Education. 2012;**43**(12):1950-1964. DOI: 10.1080/09500693.2021.1946867

[35] Kind P, Jones K, Barmby P. Developing attitudes towards science measures. International Journal of Science Education. 2007;**29**(7):871-893. DOI: 10.1080/09500690600909091

[36] Reddy L. Gender differences in attitudes to learning science in grade 7, African journal of research in mathematics. Science and Technology Education. 2017;**21**(1):26-36. DOI: 10.1080/18117295.2017.1279450

[37] Gurung GB, Moltow D, Brett P. Exploring the influence of the curriculum on student learning in culturally and ethnically diverse classroom contexts: Praxis, paradoxes and perspectives of stakeholders. Waikato Journal of Education. 2019;**24**(2):53-64. DOI: 10.15663/wje.v%vi%i.671

[38] Lehohla P. Gender series volume II: Education and gender, 2004-2014/statistics South Africa. In: Pretoria: Statistics South Africa, 2015 Report No. 03-10-12. 2015

[39] Bahar A, Adiguzel T. Analysis of factors influencing interest in STEM career: Comparison between American and Turkish high school students with high ability. Journal of STEM Education: Innovations and Research. 2016;**17**(3):64-69

[40] Petrus RM. A comparison of teachers' and students' perceptions of the factors contributing to poor performance in physical sciences: A case of South Africa. Journal of Turkish Science Education. 2018;**15**(4):93-103. DOI: 10.12973/tused.10248a

[41] Ullah R, Ullah DH, Ahmad DA. Girls academic performance in science subjects: Evidences from the industrializing and least industrialized countries. Journal of Elementary Education. 2021;**30**(2):187-200

[42] Dweck CS. Is Math a Gift? Beliefs That Put Females at Risk. American Psychological Association; 2010

[43] Changeiywo JM, Wambugu PW, Wachanga SW. Investigations of students' motivation towards learning

secondary school physics through mastery learning approach. International Journal of Science and Mathematics Education. 2011;**9**(6):1333-1350. DOI: 10.1007/s10763-010-9262-z

[44] Shepherd DL. Gender, Self-Concept and Mathematics and Science Performance of South African Grade 9 Students. 2017. Available from: https://www.ekon.sun.ac.za/wpapers/2017/wp112017

[45] Stoet G, Geary DC. The gender-equality paradox in science, technology, engineering, and mathematics education. Psychological Science. 2018;**29**(4):581-593. DOI: 10.1177 /0956797617741719

[46] Almukhambetova A, Kuzhabekova A. Factors affecting the decision of female students to enrol in undergraduate science, technology, engineering and mathematics majors in Kazakhstan. International Journal of Science Education. 2020;**42**(6):934-954. DOI: 10.1080/09500693.2020.1742948

[47] Dossi, G., Figlio, D., Giuliano, P., & Sapienza, PBorn in the family: Preferences for boys and the gender gap in math. Journal of Economic Behavior & Organization, 2021: 183, 175-188.

[48] Armstrong MA, Jovanovic J. Starting at the crossroads: Intersectional approaches to institutionally supporting underrepresented minority women STEM faculty. Journal of Women and Minorities in Science and Engineering. 2015;**21**(2):141-156

[49] Valentina L, Singh R. Gender differences in guidance needs of senior secondary school adolescents of Imphal district. International Journal of Educational Sciences. 2014;**6**(3):521-524. DOI: 10.1080/09751122.2014.11890163

[50] Clegg A. Girls and Science: A Training Module on Motivating Girls to Embark on Science and Technology Careers. 2010. Available from: https://genderit.org/resources/girls-and-science-training-module-motivating-girls-embark-science-and-technology-careers

[51] Dweck CS. Is Math a Gift? Beliefs That Put Females at Risk. American Psychological Association; 2010

[52] Osborne J, Simon S, Collins S. Attitudes towards science: A review of the literature and its implications. International Journal of Science Education. 2003;**25**(9):1049-1079

[53] Zamrud L. Exploring the motivation of female science learners towards learning science at the secondary level [Masters dissertation]. Karachi, Pakistan: Aga Khan University; 2008

[54] Terblanche H. Factors which negatively or positively affect the performance of girl learners in science. In: UNISA, ISTE Conference, October 21. 2013. p. 2013

[55] Sikhwari TD. Investigating first year students' abilities to make appropriate career choices. International Journal of Educational Science. 2015;**8**(2):363-368

[56] Mabula N. Career services provision to secondary school students in Tanzania: Is it a dream or reality? International Journal of Learning and Development. 2012;**2**:242-257

[57] Muola JM. A study of the relationship between academic achievement motivation and home environment among standard eight pupils. Educational Research Review. 2010;**5**(5):213-217

[58] Chikuvadze P, Matswetu MS. Gender stereotyping and female pupils' perception of studying advanced level sciences: A survey of one province

in Zimbabwe. Gender and Behavior. 2013;**11**(9):5285-5296

[59] Archer L, DeWitt J, Willies B. Adolescent boys' science aspirations: Masculinity, capital, and power. Journal of Research in Science Teaching. 2014;**51**(1):1-30

[60] Carrell SE, Page ME, West JE. Sex and science: How professor gender perpetuates the gender gap. The Quarterly Journal of Economics. 2010;**125**(3):1101-1144

[61] White EG. Education. Warburton Vic Signs Publishing Co.; 1933

[62] BlaŽev M, Karabegovi'c, M., Buruši'c, J., and Selimbegovi'c, L.). Predicting gender-STEM stereotyped beliefs among boys and girls from prior school achievement and interest in STEM school subjects. Social Psychology of Education. 2017;**20**:831-847. DOI: 10.1007/s11218-017-9397-7

[63] Ramsey LR. Agentic traits are associated with success in science more than communal traits. Perspectives of Individvidual Differences. 2017;**106**:6-9. DOI: 10.1016/j.paid.2016.10.017

[64] Thomas AE. Gender differences in students' physical science motivation: Are teachers' implicit cognitions another piece of the puzzle? American Educational Research Journal. 2017;**52017**(4):35-58. DOI: 10.3102/0002831216682223

[65] Šorgo A, Dojer B, Golob N, Repnik R, Repolusk S, Pesek I, et al. Opinions about STEM content and classroom experiences as predictors of upper secondary school students' career aspirations to become researcher or teachers. Journal of Research in Science Teaching. 2018;**2018**(55):1-21. DOI: 10.1002/tea.21462

[66] Sandra LH, Krywult-Albańska K. Gender andaccess to STEM education and occupations in a cross-national context with a focus on Poland. International Journal of Science Education. 2020;**42**(6):882-905. DOI: 10.1080/09500693.2020.1737341

[67] Shepherd DL. Gender, Self-Concept and Mathematics and Science Performance of South African Grade 9 Students. 2017. Available from: https://www.ekon.sun.ac.za/wpapers/2017/wp112017

[68] Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, New Jersey: Prentice-Hall, Inc; 1986

[69] Bandura A. Exercise of personal agency through the self-efficacy mechanism. In: Schwarzer R, editor. Self-Efficacy: Thought Control of Action. Washington, DC: Hemisphere; 1992. pp. 3-38

[70] Wigfield A, Eccles JS. 35 years of research on students' subjective task values and motivation: A look back and a look forward. In: Advances in Motivation Science. Vol. 7. Elsevier; 2002. pp. 161-198. DOI: 10.1016/bs.adms.2019.05.002

[71] Cohen L, Manion L, Morrison K. Research Methods in Education. Routledge; 2002

[72] Eccles J. SSubjective task value and the Eccles et al. model of achievement-related choices. In: Elliott AJ, Dweck CS, editors. Handbook of Competence and Motivation. New York, NY: Guilford Press; 2005. pp. 105-121

[73] Denzin NK, Lincoln Y. S. The discipline and practice of qualitative research. In N. Denzin and Y. Lincoln (Eds.), Handbook of Qualitative Research (4th ed.). Thousand Oaks, CA: Sage; 2011

[74] Shahid S, Bessarab D, Howat P, Thompson SC. Exploration of the beliefs and experiences of aboriginal people with cancer in Western Australia: A methodology to acknowledge cultural difference and build understanding. BMC Medical Research Methodology. 2009;**9**(1):60. DOI: 10.1186/1471-2288-9-60

[75] Petousi V, Sifaki E. Contextualising harm in the framework of research misconduct. Findings from discourse analysis of scientific publications. International Journal of Sustainable Development. 2020;**23**(3/4):149-174

[76] Buchanan D, Warwick I. First do no harm: Using 'ethical triage' to minimise causing harm when undertaking educational research among vulnerable participants. Journal of Further and Higher Education. 2021;**45**(8):1090-1103. DOI: 10.1080/0309877X.2021.1890702

[77] Shenton AK. Strategies for ensuring trustworthiness in qualitative research projects. Education for Information. 2004;**22**(2):63-75. file:///C:/Users//Trustworthinessin qualitativeresearchprojects.pdf

[78] Maguire M, Delahunt B. Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. The all Ireland journal of teaching and learning. Higher Education. 2017;**9**(3):3351-3364. http://ojs.aishe.org/index.php/aishe-j/article/view/335/553

[79] Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Medical Research Methodology. 2008;**8**(1):45. DOI: 10.1186/1471-2288-8-45

[80] Brink H, Van der Walt C, Van Rensburg G. Fundamentals of Research Methodology for Health Care Professionals. Juta and Company Ltd; 2006

[81] Ryan GW, Bernard HR. Techniques to identify themes. Field Methods. 2003;**15**:85-109

[82] Khishfe R, BouJaoude S. Lebanese students' conceptions of and attitudes towards science and related careers based on their gender and religious affiliations. International Journal of Science and Mathematics Education. 2016;**14**(1):145-167. file:///C:/KhishfeBouJaoude2016_Article_LEBANESESTUDENTS CONCEPTIONS OFA.pdf

[83] Miles MB, Huberman AM. Qualitative Data Analysis: An Expanded Sourcebook. 2nd ed. Thousand Oaks, CA: SAGE; 1994

[84] Machingambi S. Instructional strategies for motivating students: Reflections from 13 years of teaching in higher education. International Journal of Educational Sciences. 2013;**5**(3):227-235. DOI: 10.1080/09751122.2013.11890082

[85] DeWitt J, Louise AL. Who aspires to a science career? A comparison of survey responses from primary and secondary school students. International Journal of Science Education. 2015;**37**(13):2170-2192. DOI: 10.1080/09500693.2015.1071899

[86] Archer L, DeWitt J, Willis B. Adolescent boys' science aspirations: Masculinity, capital, and power. Journal of Research in Science Teaching. 2014;**51**(1):1-30. DOI: 10.1002/tea.21122

[87] Yilmaz E. Analysıs of students' success in the exam for transition to further education through some of the variables. International Journal of Academic Research Part B. 2014;**6**(1):57-63. DOI: 10.7813/2075-4124.2014/6-1/B.8

[88] Berg GVD, Coetzee LR. Academic self-concept and motivation as

predictors of academic achievement. International Journal of Educational Sciences. 2014;**6**(3):469-478. DOI: 10.1080/09751122.2014.11890158

[89] Hahn A, Banchefsky S, Park B, Judd CM. Measuring intergroup ideologies positive and negative aspects of emphasizing versus looking beyond group differences. Personality and Social Psychology Bulletin. 2015;**41**(12):1646-1664. DOI: 10.1177/0146167215607351

[90] Crasnow S, Alison W, Bauchspies WK, Potter E. Feminist perspectives on science. In: Zalta EN, editor. The Stanford Encyclopedia of Philosophy. 2015. Available from: https://plato.stanford.edu/archives/spr2018/entries/feminist-science

[91] Bandura A. Social cognitive theory in cultural context. Applied Psychology. 2002;**51**(2):269-290. DOI: 10.1111/1464-0597.00092

[92] Mghweno PE, Mghweno LR, Baguma P. Access to guidance and counseling services and its influence on students school life and career choice. African Journal of Guidance and Counselling. 2014;**1**(1):7-15. DOI: 10.1.1.674.1625&rep=rep1&type=pdf

[93] Dabula P, Makura AH. High school learners' perceptions of career guidance and development programmes for university access. International Journal of Educational Science. 2013;**5**(2):89-97. DOI: 10.1080/09751122.2013.11890065

[94] Kibirige I, Bodirwa KB. The effect of using computer simulations on grade 11 learners' performance In plants biodiversity In South Africa. Journal of Baltic Science Education. 2021;**20**(4):612-621. DOI: 10.33225/jbse/21.20.612

[95] Anders J, Henderson M, Moulton V, Sullivan A. The role of schools in explaining individuals' subject choices at age 14. Oxford Review of Education. 2018;**44**(1):75-93. DOI: 10.1080/03054985. 2018.1409973

[96] Mujtaba T, Reiss MJ. Inequality in experiences of physics education: Secondary school girls' and boys' perceptions of their physics education and intentions to continue with physics after the age of 16. International Journal of Science Education. 2013;**35**(11):1824-1845. DOI: 10.1080/09500693.2012.762699

[97] Hyde J, Mertz JM. Gender, culture, and mathematics performance. Proceedings of the National Academy of Sciences, PNAS. 2009;**106**(22):8801-8807. DOI: 10.1073/pnas.0901 265106

[98] Ramnarain UD. Understanding the influence of intrinsic and extrinsic factors on inquiry-based science education at township schools in South Africa. Journal of Research in Science Teaching. 2016;**53**(4):598-619. DOI: 10.1002/tea.21315

[99] Wise PM, Simmons ES. Improving the experience of women in the science, technology, engineering, and mathematics professoriate. In: Challenges in Higher Education Leadership: Practical and Scholarly Solutions. Taylor and Francis; 2017

[100] Kalogiannakis M, Ampartzaki M, Papadakis S, Skaraki E. Teaching natural science concepts to young children with mobile devices and hands-on activities. A case study. International Journal of Teaching and Case Studies. 2018;**9**(2):171-183. DOI: 10.1504/IJTCS.2018.090965

[101] Kalogiannakis M, Papadakis S. The use of developmentally Mobile applications for preparing pre-service

teachers to promote STEM activities in preschool classrooms. In: Papadakis S, Kalogiannakis M, editors. Mobile Learning Applications in Early Childhood Education. IGI Global; 2020. pp. 82-100. DOI: 10.4018/978-1-7998-1486-3.ch005

[102] Dorouka P, Papadakis S, Kalogiannakis M. Tablets & apps for promoting robotics, mathematics, STEM education and literacy in early childhood education. International Journal of Mobile Learning and Organisation. 2020;**14**(2):255-274

[103] Chatzopoulos A, Kalogiannakis M, Papadakis S, Papoutsidakis M, Elza D, Psycharis S. DuBot: An open-source, low-cost robot for STEM and educational robotics. In: Papadakis S, Kalogiannakis M, editors. Handbook of Research on Using Educational Robotics to Facilitate Student Learning (Pp. 441-465). IGI Global; 2021. DOI: 10.4018/978-1-7998-6717-3.ch018

[104] Vlasopoulou M, Kalogiannakis M, Sifaki E. Investigating Teachers' attitude and Behavioral intentions for the impending integration of STEM education in primary school. In: Papadakis S, Kalogiannakis M, editors. Handbook of Research on Using Education Robotics to Facilitate Student Learning. Hershey, PA: IGI Global; 2021. pp. 235-256. DOI: 10.4018/978-1-7998-6717-3.ch009

[105] Marie PE, Morales RAA, Espinosa AA. Does gender inequality influence interest in pursuing a career in science or mathematics teaching? Issues in Educational Research. 2016;**26**(1): 65-81. https://www.iier.org.au/iier26/morales.html

[106] Skorikov VB. Continuity in adolescent career preparation and its effects on adjustment. Journal of Vocational Behaviour. 2007;**70**:8-24

Chapter 6

The Power in Groups: Using Cluster Analysis to Critically Quantify Women's STEM Enrollment

Ann M. Gansemer-Topf, Ulrike Genschel, Xuan Hien Nguyen, Jasmine Sourwine and Yuchen Wang

Abstract

Despite efforts to close the gender gap in science, technology, engineering, and math (STEM), disparities still exist, especially in math intensive STEM (MISTEM) majors. Females and males receive similar academic preparation and overall, perform similarly, yet females continue to enroll in STEM majors less frequently than men. In examining academic preparation, most research considers performance measures individually, ignoring the possible interrelationships between these measures. We address this problem by using hierarchical agglomerative clustering – a statistical technique which allows for identifying groups (i.e., clusters) of students who are similar in multiple factors. We first apply this technique to readily available institutional data to determine if we could identify distinct groups. Results illustrated that it was possible to identify nine unique groups. We then examined differences in STEM enrollment by group and by gender. We found that the proportion of females differed by group, and the gap between males and females also varied by group. Overall, males enrolled in STEM at a higher proportion than females and did so regardless of the strength of their academic preparation. Our results provide a novel yet feasible approach to examining gender differences in STEM enrollment in post-secondary education.

Keywords: STEM, gender, cluster analysis, female, enrollment

1. Introduction

Gender disparities in STEM enrollment in college continue to receive a significant amount of attention [1]. Research investigating reasons for this disparity in enrollment by gender highlights three consistent themes: a) females who are less academically prepared in math and sciences than males are less likely to enroll in STEM; b) despite comparable academic preparation, females, on average, enroll in STEM majors in smaller percentages than males, and c) females' participation varies by major, with some majors, such as biology seeing a higher proportion of females than in math intensive fields (MI STEM) such as engineering or computer science [2–5]. Despite the abundance of research on this topic, there lacks a more detailed yet cohesive look at the interrelationships among high school academic preparation, ACT scores, and STEM enrollment.

This study utilized hierarchical agglomerative clustering to analyze high school and college data from a cohort (n = 3104) of students at a large, public research institution in the Midwest. Students entering this institution tend to arrive with a broad academic background and varying levels of readiness to pursue a STEM degree. Students receive guidance from professional academic advising staff and faculty about course enrollment and academic trajectories. Frequently these decisions are made by considering one or two individual data points such as a students' GPA or ACT score, yet research in this area consistently demonstrates that it is a combination of several of these factors that more accurately represents students' preparedness and ability to succeed [6–8]. By clustering, our data analysis will provide more nuanced insights than traditional statistical analyses. Traditional analyses tend to focus on averages and distributions of individual student characteristics across all students. Clustering algorithms seek to identify natural groupings (clusters) of students such that students within a cluster are academically more similar to each other regarding their pre-collegiate training than they are to students from any other cluster. Thus, clustering better accounts for the interrelationships of several factors and offers much more robust information than the more common approach to examining individual variables.

As it relates to our study, we were interested in the potential for this technique to provide a more in-depth understanding of female students' enrollment in STEM based on several factors of academic preparation. This approach has been used to study gender inequality in the STEM workforce (see, for example, [9]); we apply a similar method to examine gender disparities in STEM choice.

We begin by investigating the potential of identifying unique groups of students using hierarchical agglomerative clustering. We then examine enrollment in STEM by gender and clusters. The following questions guide our analyses:

1. What are pre-collegiate grades, academic rank, and academic mathematical and science courses, of males and females in NonSTEM, math-intensive STEM (MISTEM), and other STEM (OSTEM)?

2. Can hierarchical agglomerative clustering meaningfully identify unique groups of students based on precollege student characteristics? If yes, do gender differences in academic background exist?

3. Given cluster membership, are there gender differences in enrollment into STEM, specifically into MISTEM and OSTEM majors?

The results of our study have implications for secondary and postsecondary education. A more robust understanding of the interrelationships among variables that contribute to enrollment in STEM areas can be used to develop strategies that can enhance STEM enrollment. Postsecondary institutions can gain a better and more sophisticated understanding of the academic readiness of a cohort of incoming students. This information can be used strategically by institutions and departments to tailor course offerings to the immediate academic needs of students allowing for possibly critical adjustments in the present time. This approach can also be used to better understand not only enrollment, but also retention and completion of STEM degrees. Further, using the clusters to measure the pursuit of STEM degrees, retention, and completion of STEM degrees permits a novel use of clusters as a statistical predictive model in STEM education. Our data analysis approach is innovative and differs from traditional statistical analyses by using hierarchical agglomerative clustering that allows students with similar characteristics to be grouped together. Through the use of clustering, this study more closely examines gender differences in academic preparation, STEM interest, and enrollment in STEM.

2. Literature review

Females have been narrowing the gaps in math and science achievement and have seen more participation in STEM enrollment and careers [10]. However, the rate of participation is still drastically disproportionate, with only 27% of all STEM careers being occupied by women [11]. These statistics also vary based on major. For example, women represent over half of all bachelor's degree recipients in biology but are significantly underrepresented in physical sciences, engineering, mathematics, and computer sciences (MISTEM) [12–14]. Even at selective institutions with a large pool of interested students willing to enroll, females represent between 15 and 28% for Bachelor's degrees and only 13–20% for Ph.D.s in math departments [15]. We drew upon past literature to better understand differences in STEM and guide our selection of variables.

2.1 Pre-collegiate preparation and STEM enrollment

According to Weeden, Gelbgiser, and Morgan [16], between 19% and 32% of the gender gap for STEM degree completion can be attributed to the gender gap in STEM career interest in high school. Additionally, only 13% of female high school graduates expressed an intent to pursue a STEM career compared to 26% of their male counterparts [16]. This gap in interest as early as high school indicates that the STEM career gap is not solely caused by attrition in college or women exiting STEM careers post-graduation alone.

In high school, females consistently outperform males in their core classes, including math and science [17–19]. Despite earning high grades, females do not perform quite as well on high stakes standardized tests in math and science, scoring 0.7, 0.2, and 0.4 fewer points than males on the math, science, and STEM portions of the ACTs, respectively [20]. In spite of earning better grades in math and science, course selection among males and females shows some discrepancies. For example, female graduates took roughly the same number of advanced math courses as their male counterparts, experiencing an overrepresentation only in Algebra II. Science, however, shows more variations, with ten percent more females taking Advanced Biology, about six percent more females taking Chemistry, about 6 percent more males taking Physics, and males enrolling in engineering courses five times higher than females [21]. Furthermore, correlational research supports that participation in Advanced Placement (AP) STEM courses and STEM career interest are associated [22] and that students with high math abilities and exposure to rigorous courses were more likely to enroll in STEM majors [23].

GPA is positively correlated with the pursuit of a STEM major [24], and it is a better predictor for college success than the ACTs [25]. The rigor of math and science courses is a better predictor for enrollment in a STEM major than the number of courses alone [26]. Class rank, on the other hand, is more complex. When ranks are calculated by subject (i.e., math and reading) and communicated to students, ranking can have statistically significant effects on students' career choices. For example, a study conducted in Ireland found that students ranked highly in math had a positive association with STEM career choice and a negative association with careers in the arts and social sciences, while those who were highly ranked in English had a positive association with arts and social sciences and a negative association with STEM careers [27]. A study performed in Florida found that high school class rank and GPA, which are higher for females, were the best predictors of collegiate GPA and the number of credits earned in college [28]. But, as detailed above, males and females experience similar pre-collegiate STEM preparation in many respects with small differences in math, science, and STEM test scores and some discrepancy in enrollment of advanced science courses. Despite these similarities, males are twice as likely to intend to declare a STEM major than females. A closer look at STEM enrollment is required.

2.2 STEM enrollment

In high school, females are less likely to be interested in STEM and more likely to lose interest over time [29]. Controlling for math achievement and aptitude, females are still less likely than males to be interested in STEM [30]. In fact, one of the best predictors for enrollment and persistence through a STEM major is an individual's desire to pursue a STEM career in high school, with those expressing interest in high school completing degrees at three times the rate of those who do not express this interest in high school [31]. Among females who intend to major in a STEM field in college, nearly half of them switch majors to non-STEM fields compared with only a third of males [21].

That is not to say women are not enrolling in STEM majors; in fact, women earned 53% of STEM degrees (short of their 58% share of all degrees that would be proportional to their overall makeup of the workforce) [32]. However, there are significant disparities among the types of STEM degrees women choose to pursue. For example, women are overrepresented in health-related STEM careers with, 85% of Bachelor's degrees being awarded to women, but they are awarded less than 45% in Mathematics and Physical Science and less than 25% in Engineering and Computer Science [32]. Because interest in STEM begins before students enroll in postsecondary education and gender gaps in STEM still persist, we consider new ways to understand reasons for these gaps based on high school preparation.

3. Methods

3.1 Data source and sample

Data for this study were provided by the institution's Office of the Registrar at one large public, research-intensive institution. To ensure our research conformed to standards and guidelines of ethical research practice, we received approval from the Institutional Review Board at the study's institution. Per written agreement with the Office of the Registrar, all students enrolled in an introductory level mathematics or statistics course in the Spring 12, Fall 12, or Spring 13 semester were eligible for the parent study. Students were given the opportunity to opt out of the study, and of 16,401 eligible students, 32 chose to opt out. We focused on these courses as they often serve as gatekeeper courses to a STEM degree.

Because we were interested in the relationship between high school preparation and enrollment in a STEM major, we focused only on first-semester, degree-seeking students who entered the institution directly from high school and were enrolled in an introductory level math course. We excluded students who transferred from another post-secondary institution because we did not have access to pre-college data for many of these students. Students who were classified as non-degree-seeking or international were also removed prior to the analysis as these students are likely to differ in their academic background or degree goal. Of the initial 3219 eligible students, we had complete data on the variables of primary interest for 3104 students.

Using the STEM Designated Degree Program List (2012 revised list) provided by the Department of Homeland Security, we categorized students into STEM and NonSTEM majors. We further split STEM majors into math-intensive STEM (MISTEM) and Other STEM (OSTEM) majors. For the purpose of this study, a STEM major was considered math-intensive if it required at least one semester of science or engineering calculus. This definition is similar to the definitions used by Ceci and Williams [33] and by Bressoud [34]. This differentiation served distinct purposes: (a) the gender gap has historically been more pronounced in MISTEM majors such as engineering, computer science, or physics, whereas fields like biology or chemistry (OSTEM) have increasingly grown the proportion of women to the extent that women are now in the majority of degree earners [35]; (b) definitions of STEM vary greatly and can range from more inclusive by considering fields such as psychology, dietetics majors or kinesiology as STEM fields to less inclusive lists, which consider mainly engineering, mathematics, physics, natural sciences, and computer sciences. Distinguishing MISTEM majors define majors represented in most STEM field definitions.

Our analysis included the following variables: gender, major, high school rank (HS Rank), grade point average (GPA), number of high school credits earned in mathematics courses, including algebra, geometry, trigonometry, and calculus, and credits earned in biology, chemistry, and physics, ACT composite, ACT English, and ACT Math scores. The ACT is a national standardized test commonly used in college admissions decisions.

3.2 Statistical methodology

Using student demographic characteristics and pre-college academic background variables, we conducted an agglomerative hierarchical cluster analysis. "Cluster analysis is a data-mining technique that allows researchers to cluster a set of observations into similar (homogeneous) groupings based on a set of features" [36]. It accounts for the different high school experiences and preparation with which students enter their first year of college and can provide a more complex description of students than a comparison on a variable-by-variable basis. Clustering students reduces the focus on mean comparisons, which captures a population's average behavior, but less on how students compare at the individual level. Although cluster analysis has been used in a variety of academic settings, its use to investigate female enrollment discrepancies in STEM vs. non-STEM fields is novel. Cluster analysis has been used to develop classroom observation tools [36], reveal different learner profiles based on motivation, achievement, needs satisfaction, etc. [37], and the differences between females and males who succeed within higher technical education [38]. Using similar methods as these studies, we will compare clusters of similar students to determine any trends among factors such as gender, preparation, and STEM enrollment.

3.3 Data analysis

All statistical analyses were run in SAS/STAT software, Version 9.4 of the SAS system and RStudio Version 1.3.1073. To address the first research question, we calculated the means and standard deviations of each of the variables used in the cluster analysis by type of major (NonSTEM, MISTEM, OSTEM) and by gender. To address the second research question, we utilized PROC Cluster and Ward's minimum-variance method [39]. Ward's minimum-variance method is based on the total error sum of squares that arises by grouping observations into distinct clusters where the total sum of squares corresponds to the sum of the within-cluster sum of squares [40]. Merging a set of observations into a cluster can be considered a loss of information. Ward's method seeks to minimize the loss of information from merging any two clusters at a given step in the clustering algorithm. That is, the two clusters whose merging will lead to the smallest increase in the total error sum of squares will be combined into a new cluster [40]. Initially, each student represents a single cluster. At each step of the algorithm, two existing clusters are merged until only one cluster remains. The number of clusters is unknown prior to the analysis and an appropriate cluster solution is typically based on a set of clustering criteria such as the cubic clustering criterion (CCC), the Pseudo-F, and Pseudo T2 statistic [41, 42].

In order to see if the gender disparity in STEM enrollment is associated with gender or merely high school preparation, we clustered students according to their high school science, mathematics, and standardized test score data. Taking calculus in high school is a strong predictor of STEM interest and success [43]. Therefore, we separated students into two groups prior to clustering: students with calculus in high school (Calc group) and students without (NonCalc group). We then ran a separate cluster analysis based on high school rank, ACT English, ACT Math, and the sum of high school science credits in biology, physics, and chemistry. We decided on an initial number of clusters in each group based on the CCC, Pseudo-F, and Pseudo-T2 clustering criteria. To address the final research question, we examined the proportion of males and females in each cluster that chose a major in NonSTEM and STEM. We then limited our sample only to those who chose STEM and examined the proportion of males and females in each cluster who chose MISTEM or OSTEM.

3.4 Limitations

We wish to acknowledge some methodological limitations. To be included in the sample, a student had to take a mathematics or statistics course during their first semester in college. We are therefore missing students who may have transferred credits into college or postponed taking a mathematics or statistics course in their first semester. Our sample also only included students who chose to major in STEM in their first semester. Additional research that examines students who may decide to major in STEM after their first semester would provide additional insights.

A cluster analysis using different variables would likely result in different clusters. For example, if we were to treat the numbers of biology credits, chemistry credits, and physics credits as separable variables rather than use their sum, clusters would likely form around differences between students with respect to the individual variables such as students with many biology credits versus students with few biology credits. We chose the sum of all science credits for two reasons. From a methodology point of view, the sum of science credits is more preferable as a variable because it has a greater range of values. Secondly, the choice of variables depends on the characteristics deemed meaningful to identify differences between students.

Additionally, we wish to acknowledge the limitations and ethical considerations of using quantitatively techniques to group students and subsequent interpretations of these efforts. Quantitatively analysis affords an opportunity to see patterns that may otherwise be unclear, yet this approach can also minimize nuances within clusters and overlook significant implications of variables that were not included. For example, our study focused on gender but did not consider variables such as socioeconomic status, nationality or race, or secondary school quality. Our results also should not be used to imply causality or judgment [44, 45]– we seek to understand possible associations between variables but cannot conclude that one set of patterns causes a specific outcome or that one is qualitatively better than others.

4. Results

4.1 Research question 1. what are advanced placement scores, pre-collegiate grades, academic rank, and academic mathematics and science courses, of males and females in NonSTEM, math-intensive stem (MISTEM), and other stem (OSTEM)

Across all fields, NonSTEM, MISTEM, and OSTEM, female students are equally prepared as men in the mathematics and sciences courses (see **Table 1**). Females have consistently higher high school ranks and GPA scores compared to their male peers, which is consistent with the results of the American Association of University Women Educational Foundation [17], Degol et al., [18], and Voyer and Voyer [19]. Females who enroll in MISTEM also score on average as well as their male peers and slightly outperform them on the English ACT placement test. Men enrolling in NonSTEM and OSTEM majors show a slight advantage on the Math ACT placement test, which is consistent with prior research [20].

4.2 Research question 2: can hierarchical agglomerative clustering meaningfully identify unique groups of students based on pre-college student characteristics? If yes, do gender differences in academic background exist?

Based on the clustering criteria, four or five clusters were reasonable choices for students with and without calculus. To arrive at the most meaningful number of clusters for each group, we plotted each clustering variable using side-by-side boxplots (see **Figure 1**). Each boxplot shows the distribution of the variables for all students in the respective cluster, while the horizontal line inside the box represents the median value observed for these students. We based our decision on the final number of clusters for each group of students on what we considered to be meaningful differences in the distribution and median value for each cluster in the context of our research questions [46]. Due to the agglomerative nature of the clustering procedure, a solution consisting of four clusters arises from the merging of the two closest clusters in the solution consisting of five clusters while the remaining clusters remain unchanged. Thus, we will decide, for example, on four clusters if merging the two closest clusters in the five-cluster solution does not result in a sufficiently large loss of information but changing from four to three clusters would. In essence, we

Variable	NonSTEM NF=537, N_M=542		MISTEM NF=273, N_M=1385		OSTEM NF=216, N_M=151	
High School	Female	Male	Female	Male	Female	Male
GPA	3.62 (0.36)	3.38 (0.42)	3.81 (0.38)	3.63 (0.4)	3.65 (0.34)	3.5 (0.4)
Rank	77 (14.89)	68.1 (16.23)	83.9 (14.36)	76.6 (16.17)	77.9 (14.46)	72.3 (17.17)
Calculus Cr	0.4 (0.8)	0.5 (0.91)	1.3 (1.17)	1.2 (1.18)	0.6 (0.98)	0.5 (0.89)
Algebra Cr	4 (0.26)	4 (0.28)	4 (0.35)	4 (0.21)	4 (0.26)	4 (0.27)
Geometry Cr	2.6 (0.51)	2.6 (0.5)	2.7 (0.46)	2.8 (0.45)	2.6 (0.49)	2.6 (0.51)
Trigonometry Cr	1 (0.68)	1 (0.66)	1.2 (0.57)	1.2 (0.6)	1 (0.63)	1 (0.67)
Statistics Cr	0.3 (0.64)	0.4 (0.78)	0.4 (0.77)	0.3 (0.69)	0.3 (0.67)	0.2 (0.57)
Adv. Math Cr	1.3 (1.05)	1.2 (0.99)	1.8 (0.95)	1.7 (0.90)	1.3 (1.31)	1.3 (1.12)
Physics Cr	1 (1.03)	1.3 (1.05)	1.9 (1.09)	2 (1.05)	1.1 (0.99)	1.5 (1.01)
Biology Cr	2.7 (1.07)	2.5 (0.95)	2.6 (1.19)	2.4 (0.91)	3.1 (1.36)	2.8 (1.32)
Chemistry Cr	2.1 (0.63)	2 (0.71)	2.4 (1.03)	2.3 (0.95)	2.2 (0.69)	2.1 (0.94)
Science Cr	5.8 (1.6)	5.9 (1.47)	6.9 (2.01)	6.7 (1.66)	6.4 (1.83)	6.4 (1.92)
ACT						
Math	23.3 (3.71)	24.3 (3.68)	27.3 (4.27)	28.1 (3.89)	23.3 (4.1)	24.8 (4.12)
English	24 (4.56)	22.9 (4.28)	26.7 (5.06)	25.4 (4.6)	23.6 (4.57)	23.4 (4.8)
Composite	23.8 (3.48)	24.1 (3.47)	26.9 (4.01)	26.9 (3.63)	23.9 (3.77)	24.6 (3.62)

Table 1.
High School mathematical and science background of incoming students by enrollment into MISTEM, OSTEM and NonSTEM. All values are rounded to the nearest decimal place.

are looking for a solution that is both inclusive and parsimonious. For this reason, we included the three- and six-cluster solutions in the decision-making process. For simplicity, we discuss the different cluster solutions in terms of one cluster being split as opposed to two clusters being merged.

4.2.1 Clustering for students with calculus

We begin with the interpretation of the three-cluster solution and describe which of the three clusters is divided in the transition from three clusters to four.

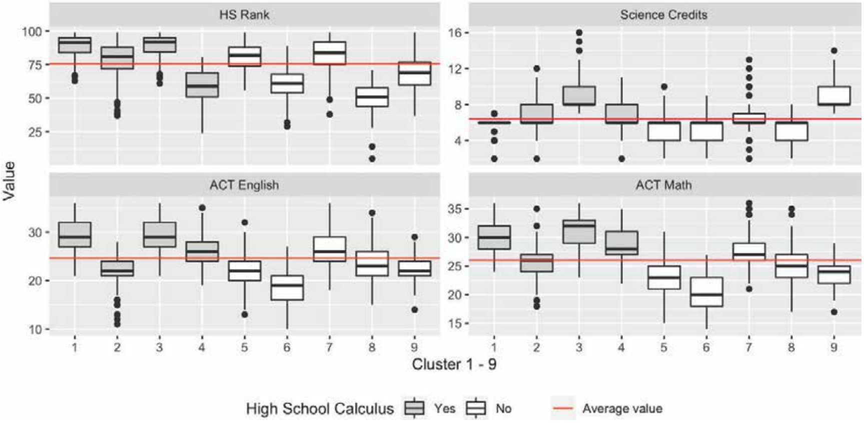


Figure 1.
Finalized cluster solution.

For students with calculus (N = 1280), a three-cluster solution identifies three types of students. A first cluster consists of 627 students with noticeably lower HS ranks, ACT English and ACT Math scores compared to the students in the remaining two clusters. Among the latter, the more academically prepared students, two groups emerge; a cluster with students who have substantially more science credits and a slightly better but noticeable ACT Math score (N = 335) compared to the students in the second group (N = 318). Both groups have comparable ACT English scores and HS Ranks.

The four-cluster solution arises from splitting the lower performing group of 627 students into two distinct groups based on differences in ACT scores and HS Ranks. Students in both groups have the lowest HS ranks out of all calculus students but the first group (N = 213) performs substantially better on the ACT English and ACT Math exams than the second group (N = 414). We deemed this split meaningful. A plausible interpretation could be that the second group (N = 414) consists of students who tend to underperform on standardized tests for a variety of reasons. Alternatively, the smaller group of students (N = 213) excels on standardized exams relative to their overall high school performance as reflected by high school rank. Due to the hierarchical nature of the clustering algorithm, the other two clusters remained unchanged.

In a five-cluster solution, the group of 414 students is broken up into students with slightly higher ACT scores, noticeably more science credits, and better HS ranks (N = 99) compared to the second group (N = 315). Although this distinction might be relevant, it was not relevant for our research questions, and we decided against this additional split; both groups of students maintained their relative, below-average performance on the ACT exams. Consequently, we determined four clusters to be the most appropriate number of clusters for students with calculus. For this reason, we did not consider the six-cluster solution.

4.2.2 Clustering students without calculus

For students without calculus (N = 1284), a three-cluster solution distinguishes three groups. Students in group 1 (N = 663) have overall the lowest high school ranks but perform otherwise similar to students in a second group (N = 423) with respect to the number of science credits taken in high school and scores on the ACT exams.

The third group (N = 738) outperforms the first and second group in the number of science credits and on the ACT components but shares HS ranks similar to those in the group of 423 students.

In a four-cluster solution, the third group of 738 students is divided into two separate groups. The first group (N = 603) includes the best students with respect to HS rank and ACT English and Math scores, but students in this cluster tend to have taken fewer science credits. The remaining students (N = 135) have the most science credits among all students without calculus. Their ACT scores and HS ranks are, however, much lower compared to the top students, and they also tend to have lower HS ranks compared to the second group (N = 423) in the three-cluster solution.

To move from four clusters to five the group of students with below average ACT scores and lower HS ranks (N = 663) are divided. This split reflects the same pattern we saw in the group of students with calculus. The 663 students are separated into a group of students (N = 258) who have overall the low high school ranks but who score better on the ACT English and Math placement tests and a second group, whose students have higher high school ranks but do not perform below average on the placement tests. Because we already saw a similar distinction among the calculus students, we consider the five-cluster solution meaningful and retain it over the four-cluster solution.

The six-cluster solution focused on the cluster previously consisting of the overall most prepared students (N = 603). These students are broken up into two clusters. The smaller of the two clusters retains the best students overall (N = 228), while students in the second cluster (N = 375) perform slightly worse than the top students, they still do better than the cluster of N = 423 students. Because both groups still outperform any of the remaining clusters overall, we decided against this additional

Cluster	Name	N	Cluster description
1	Calc. Strong, Less Science	318	Highest ACT English and HS Rank, Second highest ACT Math scores, below average (less than 6) science credits
2	Calc, Average	414	Typically, about average, showing slightly above average HS Rank and science credits but tend to fall short of average ACT English and Math scores
3	Calc, Strong Overall	335	Best students overall with all students having taken above average number of science credits
4	Calc, Low HS Rank	213	Students with far below average HS Ranks that have above average ACT English and Math scores
5	No Calc, Average	423	Above average HS Rank, very few science credits, below average ACT English and Math scores
6	No Calc, Low ACTs	405	About 50% of students have HS rank 1 standard deviation below average, lowest ACT English and Math scores
7	No Calc, Strong Overall	603	Strongest performers out of all non-calculus groups but students do not score as high as two top calculus clusters
8	No Calc, Low HS Rank	258	Lowest HS Rank, below average science credits but almost all students have ACT English and Math scores within 1 standard deviation below average
9	No Calc, Average, More Science	135	Worst performers on ACT English and Math, low HS Rank but take many science credits, credits comparable to the top cluster (Cluster 3)

Table 2.
Description of clusters and respective sample sizes.

split and retain the 5-cluster solution as the final number of clusters. A description of the finalized cluster solution is given in **Table 2**.

4.2.3 Gender distribution across clusters

Our analysis revealed that we could use clustering to find meaningful differences in groups. We then examined the proportion of female students in each cluster. Out of 3104 students, 1026 are female representing 33% of the students in the sample. Assuming that there are no systematic differences in the academic background between females and males, we expect to see about 33% of each cluster to be female students. **Figure 2** shows a visualization of the proportion of females in each cluster. In Cluster 7 (No Calculus, Low ACTs), the proportion is close to the target value of 33% with 34.3%; females are slightly below 33% in Clusters 1–3 (Calculus, Strong Less Science – Strong Overall). On the other hand, females are overrepresented in three out of the five non-calculus clusters and underrepresented in Clusters 4 (Calculus, low HS rank) and 8 (No Calculus, low HS rank). Although these students had lower HS ranks they still scored well on the ACT English and Math tests relative to students in clusters that proportionally contain more female students (Clusters 5, 6, and 9). Students in Cluster 8 (No Calculus, Strong Overall) are very similar to students in Clusters 1–3 when it comes to high school rank and performance on the ACT, except they did not have calculus in high school. The proportion of female students in a calculus cluster being average or below average shows that proportionally, fewer female students take calculus in high school (33%) than their male peers (45%).

4.3 Research question 3. Given cluster membership, are there gender differences in enrollment into STEM, specifically into MISTEM, and OSTEM majors?

In the overall sample, 48% of females chose to major in STEM and 74% of males chose to major in STEM. Of those who majored in STEM, 90% of males and 56% of females enrolled in MISTEM.

Using the cluster solution identified in **Table 2**, we examine the proportion of female students by cluster. As mentioned previously, if differences in enrollment by gender are within natural variation, we can expect about one-third of the students to be female in each cluster.

Even though the proportion of females in Clusters 1–3 and 7 are similar and close to average (see **Figure 2**), the enrollment in STEM majors is strikingly different. For example, females with a calculus background are consistently more likely than females without a calculus background to choose a STEM major. This is evident when comparing Cluster 7 to Clusters 1–4: female students with no calculus background are less likely than male students to enter STEM.

Within the same cluster, thus with similar academic background, a smaller proportion of female students enroll in STEM than male students (see **Figure 3**). When we compare across gender and clusters, we see differences in this gap. For example, for Cluster 3, there is less than a 13% difference between males and females (94% vs. 81%); however, in Cluster 7, this gap increases to 31%.

The lower enrollment rate for females is especially evident for students in the NoCalc, Strong Overall cluster. We mentioned before that this NoCalc cluster is similar in background to the calculus clusters, the only difference being that the students did not have calculus in high school. Males in this cluster, however, enroll in STEM at a similar proportion to their female peers who did have calculus.

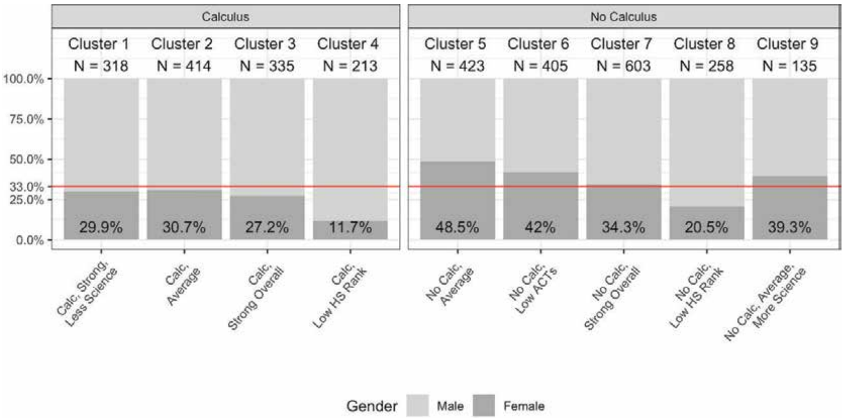


Figure 2.
Distribution of gender across clusters.

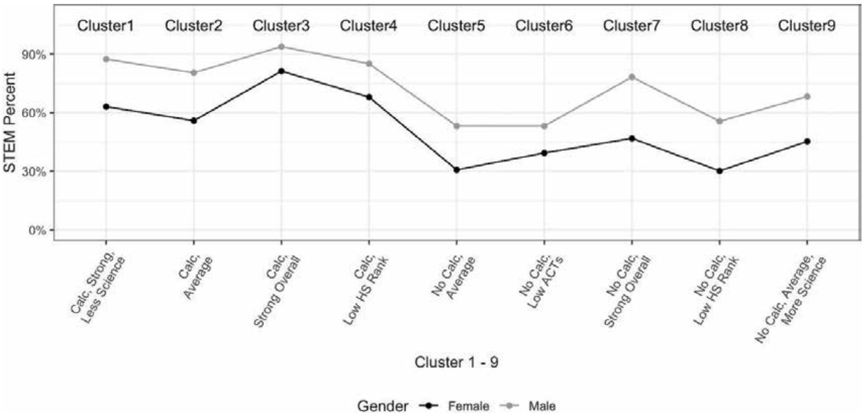


Figure 3.
Enrollment into STEM by gender.

4.3.1 MISTEM and OSTEM

We then restricted ourselves to students enrolled in STEM in each cluster. Among those, **Figure 4** shows that a higher proportion of male STEM students choose MISTEM than female STEM students. The trend is very apparent in the NoCalc clusters but is also present in the calculus clusters. Another way of saying this is that more female STEM students choose OSTEM than male students, especially the NoCalc Students. There are two interesting clusters to contrast: the NoCalc, Strong Overall and NoCalc, Average More Science. They have the same proportion of females enrolling into STEM; however, a much higher proportion of females in NoCalc, Strong Overall choose MISTEM than the NoCalc, Average More Science cluster.

Most male students in STEM, independently of their cluster membership, chose to go into MISTEM majors. Female students with a calculus background are more likely to go into MISTEM than female students without calculus in high school. The percentage of females in MISTEM is high for the NoCalc, Low ACT cluster, but because we

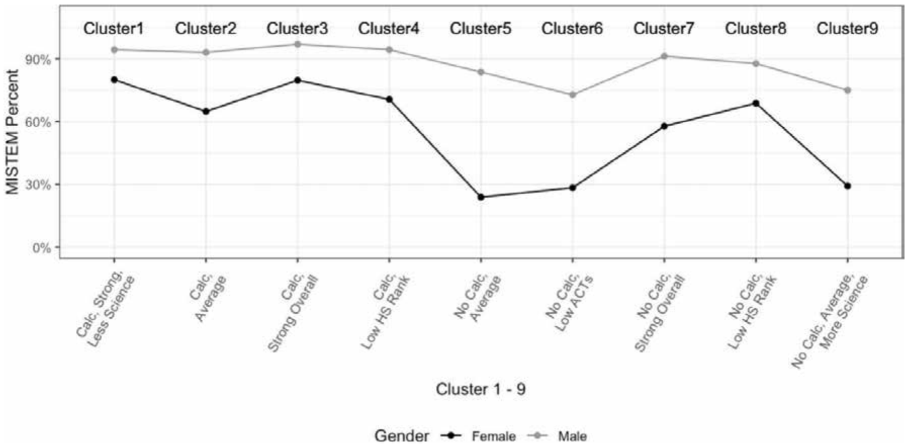


Figure 4.
Proportion of STEM students enrolled in MISTEM by gender.

limited ourselves to females who chose STEM within the cluster and the females are underrepresented in this cluster, we find an even smaller proportion of females chose STEM.

5. Discussion and implications

Despite significant efforts to minimize the gender gap in STEM, differences in interest and enrollment between men and women still exist. Ensuring a globally skilled workforce that meets the needs of the 21st century requires a post-secondary education in STEM fields, yet the interest to pursue STEM begins prior to enrolling in college. Therefore, efforts examining pre-collegiate preparation and STEM enrollment are critical.

We first examined the individual pre-collegiate variables of males and females. Similar to other research in this area [17–19], we found that overall, females and males have similar preparation although males are more likely to take calculus than females [16]. Examining these individual variables may lead to an assumption that these similarities in preparation are consistent across individuals and groups. However, by employing a more advanced statistical technique such as cluster analysis, we notice that when accounting for several pre-collegiate factors simultaneously uncovers (or reveals) marked differences in enrollment patterns within and across gender. For example, if we were to investigate just one variable, such as ACT scores, we get a different picture than when we combine variables such as standardized test scores with rank and/or GPA. For example, in our findings the combination of other factors such as enrollment in science courses, standardized test scores and high school rank results in variations in enrollment in STEM. Relatedly, the grouping of variables reveals differences in STEM participation more broadly and in the types of STEM, i.e., MISTEM and OSTEM.

Females consisted of 33% of our overall sample, yet, their representation in each cluster varies from 12% in Cluster 4 to 49% in Cluster 5 (**Figure 2**). Females never reached 33% in any of the calculus clusters. Enrollment in STEM and type of STEM (MISTEM, OSTEM) also varies by cluster. The results in **Figure 3** add/continue the pattern of female underrepresentation in STEM. Although males and females follow overall a similar trend across clusters, female enrollment in STEM is consistently below male enrollment, confirming previous results that females chose STEM at lower rates than their male peers even when they are equally prepared [3]. The consistency across clusters also confirms that this holds true for all levels of preparedness in terms of academic high school background. Patterns for males and females, however, were not entirely parallel and equal. While Cluster 3 shows the highest, and Clusters 5 and 8 the lowest enrollment in STEM for both genders differences can be found in Clusters 6 and 7. For men there is little difference between Clusters 5 and 6 and larger differences between Clusters 6 and 7. For females, these percentages increase steadily from Cluster 5 to Cluster 7. Cluster 7 tells us that, despite being strong students, females who do not have Calculus in high school are much less likely to choose STEM than male students with the same background.

Further differences exist in enrollment by type of STEM. **Figure 4** shows percentages for enrollment into MISTEM above 60% for both genders in all calculus clusters. But females in non-calculus clusters show much more variation. Although females in Clusters 5 and 8 are equally unlikely to enroll into STEM having the lowest enrollment rate overall, there is a clear difference in the rate at which females in Cluster 8 enroll into MISTEM compared to Cluster 5. Interestingly, females in Cluster 8 enroll into MISTEM in rates comparable to females in any of the calculus clusters. Female students in Clusters 5, 6 and 9 are more likely to enroll into OSTEM than MISTEM while 75% or more of male students choosing STEM still enroll into MISTEM in those clusters. Overall, male students in STEM overwhelmingly choose MISTEM ranging between 73% in Cluster 6 to 94% in Cluster 8.

Our analysis limits our ability to understand why these differences occur. However, past research may lend some insights. For example, females are more likely to possess both high verbal and high math skills, whereas males are more likely to possess solely high math skills [47]. Due to the discrepancy between math and verbal skills, males tend to choose STEM careers, whereas females, who have a choice between verbal and math-centric careers, tend to choose non-STEM-related fields, opting instead for challenging fields that are more applied and practical rather than theoretical [46]. Of course, there are also work and lifestyle factors to consider; women are looking for work-family balance and value it more highly than men [48]. In addition to lifestyle values, there are also differences between social and moral career preferences with women tending toward occupations with a social, community, or altruistic component and men tending toward careers that require working with objects [49].

Although our study cannot account for the reasons differences exist, the results have implications for research and practice. Our study illustrates a method that can be adopted by institutional leaders for use on their own campuses. Although this study was limited to one institution, we utilized commonly collected pre-collegiate data. Because of the availability of this data on most campuses, this study can be replicated in a variety of campus contexts. Institutions vary in their enrollment criteria and student populations; cluster analysis techniques afford the ability to select relevant variables and determine if unique groups emerge. Researchers have noted the importance of variables such as students' race, ethnicity, and nationality and non-cognitive variables such as self-efficacy [50]. Future research could incorporate these additional variables.

Identifying these clusters is an important first step in a more comprehensive understanding of STEM interest and success. Once established, future efforts could examine the persistence and graduation rates of students in these clusters. The research on the relationship between individual measures of academic preparation and persistence and graduation in STEM has produced mixed results. Examining the combination of these measures through cluster analysis would lead to a more nuanced understanding of the role of academic preparation in STEM enrollment. For example, if there was a consistent relationship found between completing calculus and graduation in STEM, regardless of other factors, the availability and enrollment in calculus courses in high school should be encouraged.

Qualitative research methodologies could help address the questions of "why?" Individual interviews or focus groups with students in each cluster could be conducted to understand student choices in academic preparation or what aspects of their academic preparation contributed to their enrollment and success in STEM.

6. Conclusion

Minimizing the gender gap in STEM fields continues to be necessary to meet the needs of the global workforce. Academic preparation prior to enrolling in a post-secondary institution influences students' intent to pursue STEM; yet research efforts that investigate this relationship often are limited by focusing on individual variables. Our study uses an advanced statistical technique - hierarchical agglomerative clustering - that considers multiple factors simultaneously. This technique groups students into distinct categories based on a combination of academic preparation measures and by doing so, paints a different picture of the relationship between academic preparation and STEM enrollment than simply examining individual variables. Subsequently, these inconsistencies reaffirm that narrowing the gap requires a multi-faceted approach that consider academic preparation and non-cognitive factors. In addition to its research significance, there are valuable practical implications from this work. We demonstrate how this technique can be applied to institutional data; thus, we provide a valuable tool that can be utilized in postsecondary institutions for postsecondary leaders to utilize in understanding enrollment patterns within their institutions. Summarily, our study contributes to both research and practice through its use of a robust yet accessible technique that can be widely applied to quantitative data to uncover unique patterns largely overlooked by other approaches.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. HRD 1036791.

Conflict of interest

The authors declare no conflict of interest.

Author details

Ann M. Gansemer-Topf[1]*, Ulrike Genschel[2], Xuan Hien Nguyen[3], Jasmine Sourwine[1] and Yuchen Wang[2]

1 School of Education, Iowa State University, USA

2 Department of Statistics, Iowa State University, USA

3 Department of Mathematics, Iowa State University, USA

*Address all correspondence to: anngt@iastate.edu

References

[1] National Science Board. Science and Engineering Indicators 2020 [Internet]. 2020. Available from: https://ncses.nsf.gov/pubs/nsb20201

[2] Crisp G, Nora A, Taggart A. Student characteristics, pre-college, college, and environmental factors as predictors of majoring in and earning a STEM degree: An analysis of students attending a Hispanic serving institution. American Educational Research Journal. 2009;**46**(4):924-942. DOI: 10.3102%2F0002831209349460

[3] National Center for Education Statistics. Digest of Education Statistics, Table 326.30 [Internet]. 2018. Available from: https://nces.ed.gov/programs/digest/d18/tables/dt18_326.30.asp

[4] Nix S, Perez-Felkner L, Thomas K. Perceived mathematical ability under challenge: A longitudinal perspective on sex segregation among STEM degree fields. Frontiers in Psychology. 2015;**6**:530. DOI: 10.3389/fpsyg.2015.00530

[5] Perez-Felkner L, Nix S, Thomas K. Gendered pathways: How mathematics ability beliefs shape secondary and postsecondary course and degree field choices. Frontiers in Psychology. 2017;**8**:386. DOI: 10.3389/fpsyg.2017.00386

[6] Saunders-Scott D, Braley MB, Stennes-Spidahl N. Traditional and psychological factors associated with academic success: Investigating best predictors of college retention. Motivation and Emotion. 2018;**42**:459-465. DOI: 10.1007/s11031-017-9660-4

[7] Hepworth D, Littlepage B, Hancock K. Factors influencing university student academic success. Educational Research Quarterly. 2018;**42**(1):45-61

[8] Redmond-Sanogo A, Angle J, Davis E. Kinks in the STEM pipeline: Tracking STEM graduation rates using science and mathematics performance. School Science and Mathematics. 2016;**116**(7):378-388. DOI: 10.1111/ssm.12195

[9] Olave BMT. Underestimating the gender gap? An exploratory two-step cluster analysis of STEM labor segmentation and its impact on women. Journal of Women and Minorities in Science and Engineering. 2019;**25**(1):53-74. DOI: 10.1615/JWomenMinorScienEng.2019021133

[10] Para E. Is the Gender Gap Narrowing in Science and Engineering? Soroptimist International. 2020 [Internet]. 2020. Available from: https://www.soroptimistinternational.org/is-the-gender-gap-narrowing-in-science-and-engineering/ [Accessed: January 4, 2022]

[11] Martinez A, Christnacht C. Women Are Nearly Half of U.S. Workforce but Only 27% of STEM Workers [Internet]. 2021. Available from: https://www.census.gov/library/stories/2021/01/women-making-gains-in-stem-occupations-but-still-underrepresented.html [Accessed: December 13, 2021]

[12] Dika, S. L., D'Amico, M. M. Early experiences and integration in the persistence of first-generation college students in STEM and non-STEM majors. Journal of Research in Science Teaching 2016; 53(3): 368-383. DOI: https://doi.org/10.1002/tea.21301

[13] National Science Foundation. Women, Minorities, and Persons with

Disabilities in Science and Engineering. [Data Set]. [Internet]. 2018. Available from: https://ncses.nsf.gov/pubs/nsf19304/data [Accessed: December 21, 2021]

[14] Schneider B, Milesi C, Brown K, Gutin I, Perez-Felkner L. Does the Gender Gap in STEM Majors Vary by Field and Institutional Selectivity? [Internet]. Teachers College Record. 2015. Available from: https://www.tcrecord.org/books/Content.asp?ContentID=18026. [Accessed: January 4, 2022]

[15] Glazer A. National Mathematics Survey [Internet]. 2019. Available from: https://math.mit.edu/wim/2019/03/10/national-mathematics-survey/ [Accessed: January 4, 2022]

[16] Weeden KA, Gelbgiser D, Morgan SL. Pipeline dreams: Occupational plans and gender differences in STEM major persistence and completion. Sociology of Education. 2020;**93**(4):297-314. DOI: 10.1177%2F0038040720928484

[17] Corbett C, Hill C, St. Rose A. Where the Girls Are: The Facts about Gender Equity in Education [Internet]. American Association of University Women. Available from: https://ww3.aauw.org/files/2013/02/Where-the-Girls-Are-The-Facts-About-Gender-Equity-in-Education.pdf [Accessed: March 9, 2022]

[18] Degol JL, Wang M. Te, Zhang Y, Allerton J. Do growth mindsets in math benefit females? Identifying pathways between gender, mindset, and motivation. Journal of Youth and Adolescence 2018; 47(5): 976-990. DOI: 10.1007/s10964-017-0739-8

[19] Voyer D, Voyer SD. Gender differences in scholastic achievement: A meta-analysis. Psychological Bulletin. 2014;**140**:1174-1204. DOI: 10.1037/a0036620

[20] ACT Inc. The ACT Profile Report -National [Data Set] [Interent]. 2020. Available from: https://www.act.org/content/dam/act/unsecured/documents/2020/2020-National-ACT-Profile-Report.pdf [Accessed: December 20, 2021]

[21] U.S. Department of Education, National Center for Education Statistics (NCES). Digest of Education Statistics. [Internet]. 2014. Available from: https://nces.ed.gov/programs/digest/2014menu_tables.asp

[22] Wai J, Lubinski D, Benbow CP, Steiger JH. Accomplishment in science, technology, engineering, and mathematics (STEM) and its relation to STEM educational dose: A 25-year longitudinal study. Journal of Education & Psychology. 2010;**102**:860-871. DOI: 10.1037/a0019454

[23] Bleske-Rechek A, Lubinski D, Benbow CP. Meeting the educational needs of special populations: Advanced placement's role in developing exceptional human capital. Psychological Science. 2004;**15**(4):217-224. DOI: 10.1111%2Fj.0956-7976.2004.00655.x

[24] Ware NC, Lee VE. Sex differences in choice of college science majors. American Educational Research Journal. 1988;**25**(4):593-614. DOI: 10.2307/1163131

[25] Westrick PA, Le H, Robbins SB, Radunzel JM, Schmidt FL. College performance and retention: A meta-analysis of the predictive validities of ACT® scores, high school grades, and SES. Educational Assessment. 2015;**20**(1):23-45. DOI: 10.1080/10627197.2015.997614

[26] Schneider B, Swanson CB, Riegle-Crumb C. Opportunities for learning: Course sequences and positional advantages. Social Psychology

of Education. 1998;**2**:25-53. DOI: 10.1023/A:1009601517753

[27] Delaney J, Devereux PJ. The effect of high school rank in English and math on college major choice. ESRI Working Paper No. 650.2020 [Internet]. Available from: http://aei.pitt.edu/102346/. [Accessed: January 4, 2021]

[28] Conger D, Long MC. Why are men falling behind? Gender gaps in college performance and persistence. The Annals of the American Academy of Political and Social Science. 2010;**627**(1):184-214. DOI: 10.1177%2F0002716209348751

[29] Sadler PM, Sonnert G, Hazari Z, Tai R. Stability and volatility of STEM career interest in high school: A gender study. Science Education. 2012;**96**:411-427. DOI: 10.1002/sce.21007

[30] Lubinski D, Benbow CP. Study of mathematically precocious youth after 35 years: Uncovering antecedents for math science expertise. Perspectives on Psychological Science 2006; 1: 316-345. DOI:10.1111%2Fj.1745-6916.2006.00019.x

[31] Maltese AV, Tai RH. Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among U.S. students. Science Education. 2011;**95**:877-907. DOI: 10.1002/sce.20441

[32] Fry R, Kennedy B, Funk C. Stem Jobs see Uneven Progress in Increasing Gender, Racial and Ethnic Diversity [Internet]. 2021. Available from: https://www.pewresearch.org/science/2021/04/01/stem-jobs-see-uneven-progress-in-increasing-gender-racial-and-ethnic-diversity/ [Accessed: December 20, 2021]

[33] Ceci SJ, Williams WM. Sex differences in math-intensive fields. Current Directions in Psychological Science. 2010;**19**(5):275-279. DOI: 10.1177/0963721410383241

[34] Bressoud DM. Historical reflections on teaching the fundamental theorem of integral calculus. The American Mathematical Monthly. 2011;**118**(2):99-115. DOI: 10.4169/amer.math.monthly.118.02.099

[35] Cheryan S, Ziegler SA, Montoya AK, Jiang L. Why are some STEM fields more gender balanced than others? Psychological Bulletin. 2017;**43**(1):1. DOI: 10.1037/bul0000052

[36] Denaro K, Sato B, Harlow A, Aebersold A, Verma M. Comparison of cluster analysis methodologies for characterization of classroom observation protocol for undergraduate STEM (COPUS) data. CBE Life Sciences Education. 2021;**20**(1). DOI: 10.1187/cbe.20-04-0077

[37] Ng BL, Liu WC, Wang JC. Student motivation and learning in mathematics and science: A cluster analysis. International Journal of Science and Mathematics Education. 2016;**14**(7):L1359-L1376. DOI: 10.1007/s10763-015-9654-1

[38] Engström S. Differences and similarities between female students and male students that succeed within higher technical education: Profiles emerge through the use of cluster analysis. International Journal of Technology and Design Education. 2018;**28**:239-261. DOI: 10.1007/s10798-016-9374-z

[39] SAS Institute Inc. SAS/STAT® 14.2 User's Guide. SAS Institute Inc., 2016

[40] Ward JH Jr. Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association. 1963;**58**(301):236-244. DOI: 10.1080/01621459.1963.10500845

[41] SAS/STAT® 9.2 User's Guide, Second Edition. Cary, NC: SAS Institute Inc, 2009.

[42] Sarle WS. Cubic Clustering Criterion. Technical Report A-108, SAS Institute Inc, 1983

[43] Hill C, Corbett A, Strose A. Why so Few? Women in Science, Technology, Engineering, and Mathematics [Internet]. 2010. Available from: https://files.eric.ed.gov/fulltext/ED509653.pdf

[44] Petousi V, Sifaki E. Contextualising harm in the framework of research misconduct. Findings from discourse analysis of scientific publications. International Journal of Sustainable Development. 2020;**23**(3-4):149-174

[45] Murnane RJ, Willett JB. Methods Matter: Improving Causal Inference in Educational and Social Science Research. New York, NY: Oxford University Press; 2010

[46] Hennig C. What are the true clusters? Pattern Recognition Letters. 2015;**64**:53-62. DOI: 10.1016/j.patrec.2015.04.009

[47] Wang MT, Eccles JS, Kenny S. Not lack of ability but more choice: Individual and gender differences in STEM career choice. Psychological Science. 2013;**24**:770-775. DOI: 10.1177%2F0956797612458937

[48] Wang MT, Degol JL. Gender gap in science, technology, engineering, and mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. Educational Psychology Review. 2017;**229**:119-140. DOI: 10.1007/s10648-015-9355-x

[49] Ceci SJ, Williams WM. Understanding current causes of women's underrepresentation in science. Proceedings. National Academy of Sciences. United States of America. 2011;**108**:3157-3162. DOI: 10.1073/pnas.1014871108

[50] Zeldin AL, Britner SL, Pajares F. A comparative study of the self-efficacy beliefs of successful men and women in mathematics, science, and technology careers. Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching. 2008;**45**(9):1036-1058. DOI: 10.1002/tea.20195

Chapter 7

Robots, Everlasting? A Framework for Classifying CS Educational Robots

Ishin Iwasaki, Corben Roszak, Parama Chaudhuri, Katherine LaRue and Caroline D. Hardin

Abstract

Educational robots are an exciting and growing field. While some (Lego Mindstorms, for example) have been around for decades, most are only a few years old and their durability is untested; exacerbating this are those only usable with apps, that may become suddenly unavailable. This has created a nascent but significant problem: schools investing significant time and money for educational robots with little ability to know if they will work for years or just days. Other fields in science, technology, education, and math (STEM) beyond computer science also encounter this issue as more educational robots and apps for those disciplines permeate the market. While this chapter analyzes this issue from a CS perspective, the lessons learned can be applied to other STEM areas. This chapter explores the history of the problem, documents several examples of devices that have succumbed, details the unique and specific needs of school customers, and introduces the Computer Science Risk Analysis Framework for Toys (CS RAFT) to help teachers and schools evaluate a device purchase based on a holistic understanding of device longevity. This study will also provide recommendations for CS and STEM educational robot designers.

Keywords: robotics, sustainability, tangibles, education

1. Introduction

Computer science plays a critical role in our technologically connected world, and students will get exposed to many technologies within their early years of schooling [1]. Computational thinking (CT) has been used to introduce students to computing principles within the context of the subject areas they are learning, which helps make it relevant to their understanding [2]. Robots for computer science education is a rapidly blossoming field, growing from a handful of devices pioneered by people like Michael J. Freeman in 1974 and Seymour Papert in the 1980s to over 80 different modern devices available today, many with expansive ecosystems of teacher and user resources [3, 4]. Teaching abstract concepts of CT through robotics is especially helpful as it makes the abstract concrete and children can observe the direct results of their programming commands on the robot's actions [5]. What distinguishes these robots from robots designed for tasks like manufacturing is that they are specifically designed and marketed to help teach introductory computer science concepts to K-12 students. They are typically small, cute, brightly colored, and have heavily anthropomorphic features. While there are some overlaps with robots designed for hobbyists, CS and other STEM Ed robots are oriented not around accomplishing a task for the task's sake, but in learning concepts through built-in puzzles or challenges. CS Ed robots are designed for children who might have no computer science experience and are helped by adults with limited CS experience. In a decade where significant efforts have been made to battle the stereotype of computer science and other STEM fields as intimidating, these educational robots are friendly, fun, and a great entry point for beginners [5].

CS Ed toys and robots are not only marketed for home use; they are also marketed to teachers and schools where they will reach more children and have a greater impact. Following a decade of breathless Kickstarter campaigns for CS Ed robots, a nascent problem is that some of the robots which are sold as preparing children for the future will not see the future themselves. These robots with volatile futures are likely to become unusable within a year or two of purchase. Some other robots, though, can turn out to be great investments that will help children discover joy in computational thinking for years to come.

While in some settings it is acceptable to buy an educational device that is likely to only work for a short period of time, most educators must be more judicious when selecting and purchasing these robots. The process of learning about, selecting, purchasing, and using these robots is an investment that needs to return a reasonable number of hours of use [6]. The longevity of these smart tangibles impacts more than just budgets; teachers and students invest significant time and emotions into robot initiatives [7]. Frustrating experiences do not help schools promote CS to their students, especially for those who are underrepresented in CS. Robots that cease to function—colloquially referred to as "bricked"—only contribute more e-waste to the world. As sophisticated educational technologies continue to evolve, it is important for educators, administrators, and other supporters of science education to be able to discern good investments from the bad. When faced with choosing between over 80 different devices (with more hitting the market every year), how does an educator know which is which? What questions should an educator ask when evaluating different products? What qualities should they look for to minimize the risks they are most vulnerable to? This chapter will explore the importance of long-term support in robots and will present a framework to assist educators in answering those questions.

The literature review will begin by discussing how other areas of computer science deal with similar problems of long-term support.

In part one of the findings section, examples of the problem with CS Ed robots will be explored. Part two will describe examples where a device was bought in a sealed package but was either unusable or had greatly reduced capabilities.

Part three of the findings section will introduce a proposed framework for educators to use when evaluating a potential CS Ed robot purchase and what kinds, and degrees, of risk may be incurred. An example of the framework being applied to a device is provided.

In limitations and future work, we will lay out a vision for further development of the framework so both consumers and designers of educational robots can move towards a future where robots are, if not everlasting, at least reasonably durable.

Finally, although this chapter approaches this problem through the lens of computer science, the analysis and framework can be applied to educational robots in other fields.

Issues around reliance on companion software, narrow compatibility with hardware, difficult to replace parts, and trust in companies with little history are not unique to CS Ed robots, and any educator considering purchasing an educational robot should be aware of these potential issues and apply the provided framework on a potential purchase.

2. Literature review

To motivate the need to address the problem of long-term use and stability in CS educational robots, we start by looking at the related issues as described in the literature.

2.1 Long-term support in computer science

Long-term support (LTS) for software is essential to the longevity of a product. LTS starts when software has all intended features released and is considered complete. During this phase, there will be no more additional features, and updates are limited to security and bug fixes. Security updates are necessary for internet-connected devices as flaws are identified after the product has been released and people have had time to analyze the software for defects. However, while some software is released with LTS, for most software it is not added. Software that is widely used or used by a major investor of the software's parent company is more likely to get an LTS stage.

A great implementation of LTS is the Ubuntu operating system's kernel. Ubuntu takes the approach of releasing new editions of their Linux kernel every 9 months and an LTS version of their kernel every 2 years. Ubuntu defines LTS as enterprise-focused, compatible with new hardware, and prioritizing testing over-development [8]. When a kernel is released as an LTS version, there will not be more features released and it will not use unstable packages or cutting edge software elements. The LTS is well documented for users: in the Ubuntu Wiki, there is a specific page on LTS defining exactly what LTS means to them and the standards they adhere to, and a short definition of LTS so users know that the kernel is different than their usual 9-month release. Users who download Ubuntu LTS kernels have a guarantee that the software they are downloading will be supported for 5 years.

Long-term support is not just a commitment to software updates and fixes, but also a commitment that the software will be funded and supported during that time. Funding is necessary so the software will have adequate customer support and other technical documentation available. When an LTS version is released, it is a sign that the company is confident that the product will last for the LTS period and that they are willing to fully support it for the full duration of the product lifetime. That confidence is not present in CS Ed robots, which creates the possibility that a school may invest in a product hoping for 5 years of service only for the company to abandon the product several months later.

2.2 Software Escrow

Software Escrow is a way to preserve the original source code in the case the original developer can no longer continue support for their work. The contract is between a client and developer, with the escrow agent acting as the storage for code that the developer creates. They perform checks and verification of code and ensure all code is submitted properly [9].

Escrow is an additional layer of guarantee on top of a license to use proprietary software; it is treated as an insurance policy if the licensed code is abandoned [10]. Source code is defined on each escrow contract. It normally includes source code, libraries, test data, databases, and documentation. Any quality assurance needs are also defined in the contract. Various methods could be used to assure code quality, but they need to be defined by the client. It is not typically guaranteed with the provided source code they could continue the development of the codebase.

2.3 Schools' technology budget

The most obvious problem with CS educational robots which quickly or unexpectedly break is the cost [11]. Many K-12 schools already struggle with the onerous burdens of affording technology with average spending of $18,000 per school per year for computer upgrades alone [12]. CS Ed robots are most effectively and substantially integrated with classroom instruction when there are enough for a class set (as opposed to one or two devices used in self-directed free time), but with the average per-student price of a CS Ed robot being $140, a classroom set quickly runs into the thousands of dollars. The additional context of many free, high-quality options to learn CS (such as Scratch or code.org) creates a high burden of expected usable hours for educators to justify an investment in CS Ed robots. Schools which serve low-income and underserved populations are both most likely to benefit from the engagement of CS Ed robots and the least likely to be able to afford to invest in an unreliable product.

Any investment in new technology is made with the expectation that the new technology will last for a reasonable duration, typically at least a couple of years, and, in the case of hardware and tangibles, withstand the normal wear and tear students may inflict.

Furthermore, school technology is often funded by grants or one-time funding opportunities that pay strictly for the hardware and not any of the maintenance costs. This is a potential problem for tangibles because their moving parts mean they have a higher maintenance burden, and companion software can have subscription fees. The Unruly Splat tangible, for example, requires a subscription for the Splats to be used, the app to be accessed, and other customer service perks. CS Ed robots with apps that are free initially may not remain free. Therefore, while grants are often used to kick start the purchase of new CS Ed robots, ongoing costs for maintenance of devices that fall outside the normal computer lab IT support will be an issue [13]. Furthermore, CS Ed robots have software that is not updated regularly may require a school to keep older iPads or tablets supported. If money is not set aside for maintenance, then high-risk CS tangibles would not be an option, no matter the educational benefits.

2.4 Teacher's time

Another limited resource for schools that deserves protection is a teacher's time. It takes a substantial investment of teacher time to select, learn to operate, maintain, and prepare classroom activities for CS Ed robots. It is also difficult for teachers to get "hands-on" experience with a robot before purchasing it: researching options is time-consuming, complicated, and must be done outside of school hours. Integrating new tangibles depends on many factors, such as compatibility with their past experiences, complexity of the technology, and if they can try it out before committing [14].

While most robots are sold with some guidance on how to use them in an educational setting, teachers still need to become familiar with the device and customize it for their teaching style and classroom context. If a robot is usable for a limited length of time, it is a waste of teacher time.

Wasting teacher time can have long-term ramifications. The field of instructional technology and CS education has a long history of new technologies which are heavily promoted yet do not fulfill the promises made, which can contribute to teacher skepticism about educational technology [15]. Therefore, it is important that teachers and schools can plan their tangible investments to be worth the investment over time.

2.5 App security

For the many CS Ed robots which rely on apps, the discontinuation of product support can create serious security vulnerabilities. Risks include exposing personal data of children used to create the user accounts, revealing payment methods for any "in-app purchases," giving malicious actors access to a device's camera or speakers, or even allowing a malicious actor to hack the app such that harmful content is displayed [16, 17]. It can even create legal liability for schools that have an obligation to protect student data [18].

2.6 Toys and E-waste

CS Ed robots which break or become bricked are likely to become electronic waste, or e-waste, as repair options are typically non-existent. The issue of large amounts of e-waste being generated by prematurely obsolete technology already exists. To give an idea of the scale of this problem, in 2012 the global e-waste totaled around 45.6 million metric tonnes [19]; in 2019 that value totaled over 53.6 million metric tonnes (on average 7.3 kilograms per capita) [20].

Tangibles can create e-waste in two ways. First, if the toys are irreparable, then those that break will be thrown away and new toys will be purchased to replace them. Secondly, if the tangibles are not backward compatible, or a new generation fully replaces an old version, then the older generations of the tangibles become obsolete and must be thrown away. In addition to the tangible itself, additional technology that is required that becomes obsolete can become a source of e-waste.

2.7 Emotional attachment

Finally, one of the features which help CS Ed robots do a great job of teaching beginner CS topics is how students become emotionally attached to the robots (which actually can contribute to students' success at solving coding challenges [21]). Children can become very emotionally attached to their robots [22, 23]. When the devices stop working, it can create distress [24, 25].

3. Methodology

This study uses artifact analysis as a methodology for analyzing the various design and user experience factors related to robots and smart tangibles. Artifact analysis is a process by which an artifact is analyzed to understand the design philosophy of an artifact and study its users [26]. Designers may use this process to generate pristine ideas for future designs. As an active process, artifact analysis may require additional primary or secondary research to understand the full implications of a product's design or usage [26]. Researchers can use various inquiry routes to closely analyze an artifact's various aspects, such as material, spatial features, functionality, and interactivity. In the process of analysis, researchers explore the data to tell them more about the aspect they are interested in [27]. The advantage of conducting an artifact analysis is that it does not require human research participants while yielding useful insights. The disadvantage of using this form of analysis is that it does not lend itself well to considering other factors outside the design or usability of an artifact, such as some of the questions we ask in our risk analysis frameworks like company history, product history, or legacy support.

In this study, having motivated the need for a way for educators to assess a CS Ed robot's stability, we purchased two dozen different CS Ed robots and then examined them for design features that would cause them to stop working. We used this information, in combination with consulting several educators who have worked extensively with CS Ed robots and customer product reviews to design the Computer Science Risk Analysis Framework for Toys (CS RAFT). To validate that the framework was accurate, comprehensive, and easy to use, we then used the framework to analyze six smart tangibles, after which we re-designed and improved the framework. The robots and smart tangibles that we analyze are Learning Resources Code and Go Robot Mouse Activity Set, Makeblock Codey Rocky, Makeblock CodeyBot, Wonder Cue Robot, Wonder Dash Robot, and LEGO Mindstorms EV3. We included the Code and Go Robot Mouse Activity Set as a typical example in this chapter; the analyzes of the others were omitted for space.

4. Findings

When it comes to CS education products there are little to no LTS options available. The phrases "Computer Science Education" and "Long-term support" returned no relevant papers on the first five pages of Google Scholar. In particular, no relevant research has been found about how to create general plans for LTS for CS education tangibles.

4.1 Issues identified

We identified a variety of potential ways in which a CS Ed robot may be rendered inoperable before its expected end-of-life. The largest risk is for devices that require apps or external programs. If the companion software is no longer updated, the product servers are taken offline, or the software becomes unavailable because it is removed from the relevant app store, then the robot may cease to be useful. An example is the Jibo robot: after the company was sold, owners were initially told their devices would largely stop functioning as the servers were turned off. The company was subsequently sold again, and as of the writing of this chapter, users are waiting to find out more [28]. These events can occur suddenly due to a company closing or the device's manufacturer gearing up to promote a new product.

Another way the devices may become inoperable is if the software is given a large enough update that outpaces the computing power of the hardware (tablets, laptops, etc.) that the school has access to.

Some devices have subscription models on their software, which may be affordably priced when the school purchases them but can increase to unaffordable levels.

The next most likely risk to a CS Ed robot's long-term stability is having parts break. Some are not designed for realistic durability given children's use profile or are constructed of custom plastic parts and electronics which are difficult or impossible to replace once broken. Others have parts that are small and easy to lose, especially in a classroom environment. Some robots may be designed with constant adult supervision in mind, which schools with low faculty-to-student ratios cannot afford.

On a product-by-product basis, the hardware may have short-term support in terms of limited warranties, but for the many that require supplemental software or a mobile app, there is often little mention of any software warranty or support. The Sphero device is a good example of a typical warranty: a 1-year warranty for the hardware but none for the software or its availability. Others (such as the Ozobot Evo, Wowee Coji, Sphero, and OsmoBot) were found to have a hardware warranty for only 30–90 days after purchase, but no software warranty. The Ozobot Evo goes as far as stating in section 13 of their Terms of Service that they do not guarantee any of the results they advertise on their website [29]. These types of warranties are a poor fit for schools that plan curriculum far in advance and very well might not get the devices in front of students within 30 days of purchase. Often there is no mention of a development life cycle for the apps or other accompanying software and no ability to predict when (or if) updates and bug fixes will occur.

4.2 Examples of abandoned robots

We analyzed two CS Ed robots that succumbed to some of the issues above: Codeybot and Codie. Codeybot was made by Makeblock after a successful 2016 Kickstarter (raising almost $200,00) and was extensively covered by popular media. Codie (made by Codie Labs), which was designed through several startup weekends and won awards, was featured in numerous press outlets while raising $96,306 on Indigogo [30].

4.2.1 Codeybot

Codeybot was expressly marketed to teachers, offering them a discount and promising, "Are you an educator? We want to help you bring coding to your classroom!." The authors were able to purchase a Codeybot new inbox on Ebay. It would turn on, but the Android app was no longer available; out of the two iOS apps, the one that let you manually control the features was last updated in 2016, and the one to learn to program was updated in 2018. When the company was emailed, they said they were no longer supporting the app as the device is in "End of Life." The website documentation had broken images. The company had not yet responded to a question about what the end of life process was and what support was offered to current owners. The device had sturdy construction but an embedded battery which will likely be the cause of its eventual failure.

4.2.2 Codie

The authors were not able to obtain a Codie robot despite an extensive search. However, despite the Indigogo campaign's claim that "Codie is designed with durability in mind," we were able to assess that devices that have shipped had limited functionality and are now likely no longer usable [31]. Although the Indigogo campaign claimed in

December of 2015 that the app would be available soon, the company shuttered in August of 2016. It appears that a handful of the almost 700 backers received devices that did work, but the software did not have “play or create” modes. One comment on the campaign said that they had received an email from the company where the company claimed its future was dependent on selling a lot of Codies to schools. One of the last comments on the campaign was: “Would have been nice if they would at least release the code and build specs so we could build our own.” While this is an extreme case of how CS Ed Robots can be negative experiences for educators, most CS Ed robots sold today are vulnerable to the issues which doomed the Codie.

4.3 CS RAFT (Computer Science Risk Analysis Framework for Toys)

We developed CS RAFT in order to assist educators to identify the risks associated with buying educational tangibles for their classrooms. The framework is divided into four major categories: Contents, Concepts, Company, and Community.

Each educator may have access to different resources, and risks that can be taken on by one educator can vary from those that can be taken on by another. As such, the framework does not aim to assign a single value for risk or make any recommendations regarding purchases; instead, it is designed to present an educator with the essential questions to answer in order to make an informed decision about the risks they are willing to take.

As an example, we analyzed the Code and Go Robot Mouse Activity Set from Learning Resources. This standalone robot does not require any additional software or hardware and therefore has no risk of software incompatibility or deprecation issues. Learning Resources is a well-established company with many products and has a low risk of dissolution. However, community feedback suggests issues with quality control on the wheels, especially over repeated use in a classroom, and a broken wheel requires the entire robot to be replaced. An educator on a grant-funded program that can afford many replacements for a shorter-term project may be willing to take on this risk, while another educator looking for a long-term toy to use as a supplement in their classroom may not.

4.4 Advice for designers

For new robots that are being developed, there should be considerations for the longevity of the device. New devices could be improved if there were plans for both software and hardware flexibility so that in the case the project is discontinued, there would still be a way for users to continue enjoying the device to its fullest potential. Reaching out to the target community and getting feedback from end-users could help generate a solid customer base for the robot to ensure the success of the product. Finally, we encourage CS Ed robot manufacturers to work towards standards for communicating with the devices such that software can be replaced if it becomes unavailable.

In addition, robots should be made as modular as possible, and replacement modules and parts should be available for purchase. As students use the robots in the classroom, not all robots will break in the same fashion. Rather than forcing educators to purchase an entirely new robot because of one broken wheel, giving them the option to replace just the wheel (perhaps with one 3D printed from provided specifications) and repair the robot themselves would be ideal. This will allow for less e-waste and long-term use in the classroom and ultimately more educators being willing to invest in devices.

5. Contributions

This research highlights and documents a new and growing problem facing K-12 CS educators: the unpredictable longevity of CS Ed robots. To address this problem, this chapter presents a novel framework, CS RAFT (Computer Science Risk Analysis Framework for Toys), which can aid educators in understanding the risk profile of CS Ed Robots when selecting them for purchase, and can be of use to robot designers and manufactures in developing more devices.

The CS RAFT framework can be applied to educational robots and toys outside of computer science. Tangibles in all facets of education are evolving to integrate more technologies like robots with complex microcontrollers, companion software, and augmented or virtual reality systems. Happy Atoms teaches students chemistry and physics through modeling atoms and compounds using plastic pieces that can then be scanned with a tablet running a companion application. The application offers more information on the scanned molecules, guided learning tools, and more. Miko 3 is an artificial intelligence-driven robot that engages kids with a variety of STEAM (Science, Technology, Engineering, Arts, and Math) topics through its catalog of games and tools. Tools like Happy Atoms and Miko 3 have similar concerns to the CS Ed robots discussed in this chapter and should be analyzed through the CS RAFT framework. Educators may ask themselves about the usability of the Happy Atoms models if the companion app stops being supported, or the cost incurred at a school of recurrent payments for the Miko Max premium subscription.

6. Conclusion

In conclusion, there are numerous and significant impacts of CS Ed robots becoming prematurely unusable, but as we find in other areas of computer science, software and hardware longevity is a problem that has a number of solutions. This problem extends to all other “smart” devices which educators purchase, and while there are legal frameworks available in some countries to give customers remediation, for now, there is little protection for consumers in the United States. Educators would benefit from being able to make purchases based on a more informed risk profile of CS Ed robots, and designers could do more to prevent this problem. The CS-RAFT framework presented in this chapter can assist anyone poised to make a significant investment in a CS Ed robot to look more critically at the qualities of the robot. Applying the framework can better inform the purchaser on the potential risks of investment and make smarter decisions about the choice, use, or quantity of the robot. We hope this work will motivate more research in this area and move us towards a future where CS Ed robots are confidently purchased by educators with the expectation that they will last long enough to validate their investment.

7. Limitations and future work

One limitation is the rapidly evolving space of how robots are used to teach computer science, of legal frameworks around “fitness for sale,” and of norms and standards in manufacturing robots. It is possible, although unlikely, that a solution will naturally arise in one of those spaces between the writing of this chapter and its publication. Another limitation is the limited amount of user testing our framework could receive from teachers who use CS Ed robots; this was primarily caused by the pandemic both reducing the use of tangibles and its impact on teachers' available time to assist in research. In future work, we will partner with teachers to extensively use and give feedback for the next iteration of the CS RAFT.

Appendix A

A.1 Risk analysis framework

Use the framework below to guide your computer science toy risk analysis, and consider your organization's needs and available resources. A toy that has high-risk factors in one category may still be worth investment. For example, a toy produced by a new startup with little history may seem risky from a company perspective, but if the toy can be easily maintained by the user, long-term company support may not be necessary.

A.1.1 Contents

The contents category considers the hardware and software of the toy itself to assess the impact of long-term wear, lost or broken materials, and discontinued support. While self-contained, independent toys run the risk of being difficult to repair or replace, they are also immune to issues arising from the interfacing tablet or computer. Educators should consider the resources available to them when evaluating the weight of the focus points in this section.

A.1.1.1 Software flexibility

- What app or software does this toy need?
- This app or software must be online and/or connected to a server. ☐
- This app or software is open-source. ☐
- This app or software is third-party software (e.g. Arduino IDE). ☐

A.1.1.2 Hardware flexibility

- Additional technologies required to play with this toy:
  - Personal computer/laptop ☐
  - Tablet (e.g., iPad) ☐
  - Other (e.g., Infrared camera, light sensor) ☐
- List any specific requirements for the hardware required (e.g., Must be an iPad; no Android tablets)
- Batteries are easily replaceable ☐

A.1.1.3 At-risk parts

Wheels, buttons, and small supplemental parts are at high risk of being lost, broken, or worn down even under proper use.

- This toy has:
    - Wheels ☐
        a. How vital are the wheels to learning with the toy?
    - Buttons ☐
        a. How vital are the buttons to learning with the toy?
    - Supplemental parts ☐
        a. List any supplemental parts that are vital for learning with the toy.

A.1.1.4 Replaceability

Consider the at-risk parts checked in the section above.

- How much does this toy cost (per unit)?
- Does the company offer replacements for at-risk parts?
- Can any of the at-risk parts be substituted by another off-the-shelf item?
- Can any of the at-risk seem repairable by the user?
- What is the warranty on the product?

A.1.2 Concepts

The concepts category considers the learning objectives, overall design of the toy, and the nature of play while using the toy. Adopting a new robot into the classroom environment requires investment beyond the price of the toy, such as lesson planning and teacher training. Choosing products that share similar concepts with others on the market can mitigate some of the risks involved in those investments, as those similar toys can be interchanged with minimal edits to the lesson plan. While there may be benefits to using a highly specialized toy, purchasing robots that are extremely unique in their conceptual design pose a higher risk to the educator of having non-transferable content if support for the toy is discontinued.

A.1.2.1 Core concepts

- What computer science concept does this toy aim to teach?
- Is this toy a self-guided toy, or does it require teacher facilitation?

A.1.2.2 Interface and accessibility

- How do people interact with the toy (e.g., touch screen, voice commands)?
- Does it meet the accessibility needs of your audience?
- How much space and preparation does this toy need?
- How many people can share the toy at once while being effective?

A.1.2.3 Estimated effective playtime

- How much time does the toy need to be played with for fun and effective learning?
- This toy requires continued, long-term progress on the same toy. ☐

A.1.3 Company

The company category considers the history and nature of the company that produces the toy. While it is difficult to predict the future of a company's direction, trends within the company for previous versions of the toy or related toys can be useful for understanding the risk of committing to using the toy long-term. Reliable support from a well-established company can mitigate some of the uncertainty involved around buying specialized, delicate, or otherwise risk-carrying toys.

A.1.3.1 Company history

- When was this company established?
- This company makes products beyond computer science education toys. ☐

A.1.3.2 Product history

- When was this product released?
- When was the last time the software for the toy received an update?

A.1.3.3 Versions and variants

- Are there previous versions or variant versions (e.g., same toy, but branded with a popular movie franchise) of the toy?
- Are previous versions still supported and marketed?
- Are parts backward-compatible with previous versions?

A.1.3.4 Legacy support

- What is the company policy on supporting legacy products?

- Is there evidence of products becoming open-source once discontinued?

A.1.4 Community

The community category considers the relationship between the toy and its various users, and analysis in this section can inform the answers to other categories in the framework. Ratings and feedback from the community can be litmus tests for company responsiveness, and they can also inform the user of risks of breakage or loss. A rich community of educators supported by first and third-party resources such as lesson plans, video demonstrations, and professional development opportunities can suggest signs of a stable toy.

A.1.4.1 User feedback

- What are the ratings and reviews from other users?
- Record any common reasons for critical comments and concerns.
    - The company responds to these public comments. ☐

A.1.4.2 Community forums

- A public forum for this toy exists (official or otherwise) for educators to participate in.
    - This community is active (most recent activity is within the last month). ☐

A.1.4.3 First-party resources

- The company offers the following resources for educators:
    - FAQs ☐
    - Educator guides/lesson plans ☐
    - Educator training/professional development ☐
    - Video guides and demonstrations ☐
    - Other resources (list them):

A.1.4.4 Third-party resources

- There are the following resources created by other educators and users for this toy:
    - Educator guides/lesson plans ☐
    - Customized extensions or parts ☐
    - Video guides and demonstrations ☐

- Student artifacts ☐
- Other resources (list them):

Author details

Ishin Iwasaki[1†], Corben Roszak[1†], Parama Chaudhuri[2†], Katherine LaRue[1†]
and Caroline D. Hardin[1*]

1 Western Washington University, Bellingham, USA

2 Indiana University Bloomington, Bloomington, USA

*Address all correspondence to: hardinc3@wwu.edu

† These authors contributed equally.

IntechOpen

References

[1] Livingstone S, Haddon L. EU Kids Online: Final Report. London, UK: London School of Economics and Political Science; 2009

[2] Yadav A, Hong H, Stephenson C. Computational thinking for all: Pedagogical approaches to embedding 21st century problem solving in K-12 classrooms. TechTrends. 2016;**60**(6): 565-568

[3] Freeman MJ. Advanced verbal computers in education. Educational Technology. 1975;**15**(5):58-60

[4] Resnick M, Ocko S, Papert S. Lego, logo, and design. Children's Environments Quarterly. Cambridge, MA: Harvard University Press's; 1988; 5(4):14-18

[5] Sullivan A, Bers MU. Dancing robots: Integrating art, music, and robotics in Singapore's early childhood centers. International Journal of Technology and Design Education. 2018;**28**(2):325-346

[6] Papadakis S, Vaiopoulou J, Sifaki E, Stamovlasis D, Kalogiannakis M, Vassilakis K. Factors That Hinder in-Service Teachers from Incorporating Educational Robotics into Their Daily or Future Teaching Practice. In: Proceedings of the 13th International Conference on Computer Supported Education. Virtual Event; 2021. p. 55-63

[7] Papadakis S, Vaiopoulou J, Sifaki E, Stamovlasis D, Kalogiannakis M. Attitudes towards the use of educational robotics: Exploring pre-service and in-service early childhood teacher profiles. Education Sciences. 2021;**11**(5):204

[8] Mahnke P. LTS—Ubuntu Wiki. Available from: https://wiki.ubuntu.com/LTS

[9] Weigl E, Binder J, Strodl S, Kolany B, Draws D, Rauber A. Framework for Automated Verification in Software Escrow, Proceedings of the 10th International Conference on Preservation of Digital Objects; 2013. pp. 95-103

[10] Gauberti A. Do you need to put in place an escrow agreement with respect to the software and/or source code that you license? 2020. Available from: https://crefovi.com/articles/do-you-need-to-put-in-place-an-escrow-agreement/

[11] Soule T, Heckendorn R. Cotsbots: computationally powerful, low-cost robots for computer science curriculums. Journal of Computing Sciences in Colleges. Oct 2011;**27**(1):180-187

[12] Ireh M. Budgeting and Funding School Technology: Essential Considerations. School Business Affairs. 76(7):18

[13] Bielefeldt T. Systemic Planning for Technology. Oregon School Study Council Bulletin. Jan 1997;**40**(2):1-32

[14] Dias LB. Integrating technology; Learning and Leading with Technology. 1999;**27**:10-13

[15] Cuban L. Oversold and Underused. Harvard University Press; 2009

[16] Yong S, Lindskog D, Ruhl R, Zavarsky P. Risk Mitigation Strategies for Mobile Wi-Fi Robot Toys from Online Pedophiles. In: 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing. Boston, MA; 2011. p. 1220-1223

[17] Albuquerque OP, Fantinato M, Kelner J, de Albuquerque AP. Privacy in smart toys: Risks and proposed solutions. Electronic Commerce Research and Applications. 2020;**39**:100922

[18] Reyes I, Wijesekera P, Reardon J, Elazari Bar On A, Razaghpanah A, Vallina-Rodriguez N, et al. "Won't Somebody Think of the Children?" Examining COPPA Compliance at Scale. 2018. Available from: https://dspace.networks.imdea.org/handle/20.500.12761/551 [Accepted: 13 July 2021]

[19] Perkins DN, Brune Drisse MN, Nxele T, Sly PD. E-waste: A global hazard. Annals of Global Health. 2014; **80**(4):286

[20] Forti V, Baldé CP, Kuehr R, Bel G. The Global E-waste Monitor 2020. Bonn, Geneva and Rotterdam: United Nations University/United Nations Institute for Training and Research, International Telecommunication Union, and International Solid Waste Association; 2020. 120 p.

[21] You S, Robert L. Emotional Attachment, Performance, and Viability in Teams Collaborating with Embodied Physical Action (EPA) Robots. Rochester, NY: Social Science Research Network; 2017

[22] Weiss A, Wurhofer D, Tscheligi M. "I love this dog"—Children's emotional attachment to the robotic dog AIBO. International Journal of Social Robotics. 2009;**1**(3):243-248

[23] Huang L, Picart J, Gillan D. Toward a generalized model of human emotional attachment. Theoretical Issues in Ergonomics Science. 2021;**22**(2):178-199

[24] Gjersoe NL, Hall EL, Hood B. Children attribute mental lives to toys when they are emotionally attached to them. Cognitive Development. 2015;**34**: 28-38

[25] Carter EJ, Reig S, Tan XZ, Laput G, Rosenthal S, Steinfeld A. Death of a robot: Social media reactions and language usage when a robot stops operating. In: Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction. New York, NY, USA: Association for Computing Machinery; 2020. pp. 589-597. DOI: 10.1145/3319502.3374794

[26] Noah L. Artifact Analysis. Available from: http://dlrtoolkit.github.io/artifact-analysis/

[27] Given LM. Artifact analysis. In: The SAGE Encyclopedia of Qualitative Research Methods. Thousand Oaks: SAGE Publications. 2008. Available from: http://www.yanchukvladimir.com/docs/Library/Sage%20Encyclopedia%20of%20Qualitative%20Research%20Methods-%202008.pdf

[28] Bartneck C. Why do all social robots fail in the market? 2020. Available from: https://ir.canterbury.ac.nz/handle/10092/101172 [Accepted: 26 October 2020]

[29] Ozobot Terms of Use. Available from: https://ozobot.com/terms-of-use

[30] Introduction to Programming—mBlockly for Codeybot. Available from: https://www.makeblock.com/project/introduction-to-programming-mblockly-for-codeybot

[31] Codie—Cute Personal Robot That Makes Coding Fun. Available from: http://www.indiegogo.com/projects/1006485/fblk

Edited by Michail Kalogiannakis
and Maria Ampartzaki

Along with focusing on advances that bring forward STEM education, this book presents research on a variety of issues related to the STEM approach. Chapters discuss strategies in STEM teaching and learning as well as strategies that help students cope with the challenges of hard mathematical work. They also address problems of resource purchasing and factors that impact girls' decisions to pursue STEM careers. Through these and other topics, this book seeks to facilitate a deeper understanding of STEM education and its challenges in the contemporary world.

Published in London, UK

IntechOpen

ISBN 978-1-80355-731-1
9 781803 557311