Research Topics

Innovations and technologies in science/STEM education: Opportunities, challenges and sustainable practices

Edited by
Wang-Kin Chiu, Hon-Ming Lam and Morris Siu Yung Jong

Published in
Frontiers in Education

ISSN 1664-8714
ISBN 978-2-8325-4702-1
DOI 10.3389/978-2-8325-4702-1

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Innovations and technologies in science/STEM education: Opportunities, challenges and sustainable practices

Topic editors
Wang-Kin Chiu — The Hong Kong Polytechnic University, China
Hon-Ming Lam — The Chinese University of Hong Kong, China
Morris Siu Yung Jong — The Chinese University of Hong Kong, China

Citation
Chiu, W.-K., Lam, H.-M., Jong, M. S. Y., eds. (2024). *Innovations and technologies in science/STEM education: Opportunities, challenges and sustainable practices*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-4702-1

Table of contents

TYPE Editorial
PUBLISHED 21 March 2024
DOI 10.3389/feduc.2024.1387540

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Lianghuo Fan,
East China Normal University, China

*CORRESPONDENCE
Wang-Kin Chiu
✉ oscar.chiu@cpce-polyu.edu.hk

RECEIVED 18 February 2024
ACCEPTED 04 March 2024
PUBLISHED 21 March 2024

CITATION
Chiu W-K, Lam H-M and Jong MS-Y (2024) Editorial: Innovations and technologies in science/STEM education: opportunities, challenges and sustainable practices. *Front. Educ.* 9:1387540. doi: 10.3389/feduc.2024.1387540

Editorial: Innovations and technologies in science/STEM education: opportunities, challenges and sustainable practices

Wang-Kin Chiu[1]*, Hon-Ming Lam[2] and Morris Siu-Yung Jong[3]

[1]Division of Science, Engineering and Health Studies, College of Professional and Continuing Education, The Hong Kong Polytechnic University, Kowloon, Hong Kong SAR, China, [2]School of Life Sciences and Centre for Soybean Research of the State Key Laboratory of Agrobiotechnology, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, China, [3]Department of Curriculum and Instruction and Centre for Learning Sciences and Technologies, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, China

KEYWORDS

science education, STEM education, education for sustainable development, teacher competency, laboratory learning, pedagogy, innovation, technology

Editorial on the Research Topic
Innovations and technologies in science/STEM education: opportunities, challenges and sustainable practices

1 Introduction

We are witnessing an unprecedented paradigm shift in the contemporary education system entering a new era of digitalization and artificial intelligence (Dai et al., 2023; Jong, 2023). Recently, the COVID-19 pandemic has further institutionalized the applications of emerging technologies and heightened their roles, accenting the new normal of embracing innovations and harnessing technologies in education (Huang et al., 2022a). The overall picture has revealed novel opportunities for teaching and learning which were not previously evident. This is particularly true in the contexts of science/STEM education where laboratory activities, as well as graphical representations and visualizations of scientific theories, are fundamental and detrimental to the facilitation of teaching and learning (Chiu, 2021; Thees et al., 2021; Wong et al., 2021).

Despite the remarkable advancement of technologies and their emerging applications in teaching and learning, the accelerating tendency of technologization has presented challenges in education (Zhai et al., 2021; Lau and Jong, 2023). For example, effective use of innovative tools and technologies would require educators to genuinely understand the significance of human connection and interaction in classroom teaching. Pedagogic design also demands teachers' professional development of competency and capability to integrate knowledge of different STEM disciplines (Jong et al., 2021; Huang et al., 2022b). Meanwhile, further work is also needed to address various issues, such as promotion of inclusion and equity, development of teaching resources for sustainability, and redefining the roles of teachers and students in science/STEM education supported by emerging technologies. All these opportunities and challenges prompt us to initiate the present collection of Research Topic.

2 Emerging themes of the Research Topic collection

This Research Topic was established to collect quality studies related to applications of emerging technologies in science/STEM education, as well as innovations in the teaching and learning of various STEM subjects. Overall, 12 articles contributed by scholars across various countries have been published; the studies involve research review, empirical research, as well as curricular, instructional and/or pedagogical design and implementation. Two prominent themes permeating the included articles are identified: (1) impact of innovations and technologies on students' learning; and (2) teacher competencies and student skills for sustainable development. The sub-sections below present an overview of the included articles.

2.1 Impact of innovations and technologies on students' learning

Wick et al. present the use of a binary classification model as a strategy to identify and help under-prepared engineering students in early foundation STEM courses. Student performance data were used to design the model and the interventions were found to achieve an overall improvement for the high-risk engineering students in terms of success and retention rates. Firetto et al. provide another example of innovations and explore the implementation of an online module regarding effective study strategies and its association with better exam performance in an introductory anatomy and physiology subject class. Besides student success, the impact of technologies on students' interest and learning experience in science/STEM education is another high-priority topic in the research agenda in the field. Tablatin et al. investigate the use of Minecraft to cultivate Filipino students' STEM interest and the effects of deploying game-based activities on students' learning experience. Gopabala Krishnan et al. highlight the potential of a graphical user interface tool in simplifying the learning process for students in STEM disciplines. Oss Boll et al. report on the design and development of synthetic biological circuits by 3D printing. The novel STEM educational resources demonstrate its usefulness as a teaching tool to facilitate students' learning and understanding of synthetic biology, a relatively new discipline in science. The work of Doore et al. provides another example of harnessing emerging technologies for graphical representations in STEM education. They report on the design and evaluation of a universally accessible multimodal system for both students with and without visual impairment, in the communication and interpretation of graphical representations.

2.2 Teacher competencies and student skills for sustainable development

Teachers' professional development and competencies, as well as students' essential skills such as digital literacy and self-organization of knowledge, have influential contributions to the sustainable development of education in this digital and intelligence era. Thyssen and Meier report on a comprehensive analysis of teachers' perceptions toward the use of 3D printing technology in classroom teaching in Germany, with regards to skill development and didactic integration of the technology in subject lessons. Li presents the pedagogical frameworks, development and implementation of open education resources to address the issues of inadequate sense of STEM identity among students when teaching foundation computational social science incorporated with data science methodologies and social science theories. Halonen et al. describe the use of AI-directed speech recognition technology in a science education context and analyze its role in the co-construction process in terms of self-organization of knowledge. The work of Liu et al. underscores the need for cross-cultural collaboration and the prominence of innovation mindset for students to address global issues in the future. Their study examines maker education through thematic analysis and identifies key educational themes such as transdisciplinary creativity and skills in relation to sustainability.

In view of the evolving landscape of education with the emergence of new technologies and innovations, periodic review and analysis are necessary and serve as important resources to identify emerging trends and research directions. Muilwijk and Lazonder report on the systematic review of studies comparing virtual and physical investigations, and the relevant findings from the meta-analysis regarding the implications on STEM teachers in the option for virtual or physical investigations. Chakraborty et al. present a bibliometric analysis of prior research attempts in the contexts of key components such as learning methodologies and competencies for the mapping between Industry 4.0 and Education 4.0.

3 Conclusion

This Research Topic collection underscores the roles of technologies and innovations in shaping the future of science/STEM education. The diversified perspectives presented in the included studies highlight the importance of continuous efforts and more research investigations to harness innovations and emerging technologies within the contexts of science/STEM education. We hope this Research Topic collection will provide a pathway for ongoing dialogue and further insights regarding the opportunities and challenges of sustainable development in education in this new era, and provoke thoughts leading to future studies of education supported by the emerging technologies and innovations for our next generations to address the global challenges.

Author contributions

W-KC: Writing – review & editing, Writing – original draft. H-ML: Writing – review & editing, Writing – original draft. MJ: Writing – review & editing, Writing – original draft.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

We thank all the scholars who have contributed to the review process and given valuable comments to the authors.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Chiu, W. K. (2021). Pedagogy of emerging technologies in chemical education during the era of digitalization and artificial intelligence: A systematic review. *Educational Sciences* 11, 709. doi: 10.3390/educsci11110709

Dai, Y., Liu, A., Qin, J., Guo, Y., Jong, M. S. Y., Chai, C. S., et al. (2023). Collaborative construction of artificial intelligence curriculum in primary schools. *Journal of Engineering Education* 112 23–42. doi: 10.1002/jee.20503

Huang, B., Jong, M. S. Y., and Chai, C. S. (2022a). Design and implementation of a video-facilitated transdisciplinary STEM curriculum in the context of COVID-19 pandemic. *Educational Technology & Society*. 25, 108–123.

Huang, B., Jong, M. S. Y., Tu, Y. F., Hwang, G. J., Chai, C. S., and Jiang, M. Y. C. (2022b). Trends and exemplary practices of STEM teacher professional development programs in K-12 contexts: A systematic review of empirical studies. *Computers & Education*. 189, 104577. doi: 10.1016/j.compedu.2022.104577

Jong, M. S. Y. (2023). Pedagogical adoption of SVVR in formal education: Design-based research on the development of teacher-facilitated tactics for supporting immersive and interactive virtual inquiry fieldwork-based learning. *Computers & Education*. 207, 104921. doi: 10.1016/j.compedu.2023.104921

Jong, M. S. Y., Song, Y., Soloway, E., and Norris, C. (2021). Teacher Professional Development in STEM Education. *Educational Technology & Society* 24, 81–85.

Lau, W. F. F., and Jong, M. S. Y. (2023). Typology of teachers' stages of concern for STEM education. *Research in Science and Technological Education*. 41, 1560–1578. doi: 10.1080/02635143.2022.2064447

Thees, M., Kapp, S., Altmeyer, K., Malone, S., Brünken, R., and Kuhn, J. (2021). Comparing two subjective rating scales assessing cognitive load during technology-enhanced STEM laboratory courses. Front. Educ.. 6, 705551. doi: 10.3389/feduc.2021.705551

Wong, C. H. S., Tsang, K. C. K., and Chiu, W. K. (2021). Using augmented reality as a powerful and innovative technology to increase enthusiasm and enhance student learning in higher education chemistry courses. *J. Chem. Educ.*, 98, 3476–3485. doi: 10.1021/acs.jchemed.0c01029

Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., et al. (2021). A Review of Artificial Intelligence (AI) in Education from 2010 to 2020. *Complexity* 8812542. doi: 10.1155/2021/8812542

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 09 December 2022
DOI 10.3389/feduc.2022.1087064

Check for updates

Measuring the impact of student success retention initiatives for engineering students at a private research university

David P. Wick[1,2*†], Emmanuel Atindama[3*†], Prashant Athavale[2,3*], Sumona Mondal[3], Michael W. Ramsdell[2,4], Robert P. Jaspersohn[4] and John C. Moosbrugger[2,5]

[1]School of Individualized Studies, Rochester Institute of Technology, Rochester, NY, United States, [2]Institute for STEM Education, Clarkson University, Potsdam, NY, United States, [3]Department of Mathematics, Clarkson University, Potsdam, NY, United States, [4]Department of Physics, Clarkson University, Potsdam, NY, United States, [5]Department of Mechanical and Aerospace Engineering, Clarkson University, Potsdam, NY, United States

Introduction: Student success in Science, Technology, Engineering, and Mathematics (STEM) is a national concern. To increase engineering retention and graduation rates at a small private institution, a university council developed a binary classifier to identify high-risk students and proposed interventions that included decoupling first-year Physics and Calculus courses, support in introductory Calculus, and Spatial Visualization (SV) training. This paper aims to validate the binary classifier used to identify the under-prepared students entering their first year and assess the impact of the interventions. We provide a comparative analysis of student success metrics for high-risk engineering students across a decade of cohorts, including 5 years before (2006–2010) and 5 years after (2011–2015) implementation of intentional strategies.

Methods: We validated the binary classifier using an accuracy measure and Matthews Correlation Coefficient (MCC). We used the 2-population proportion test to compare STEM retention and 4- and 6-year graduation rates of High-Risk engineering students before and after interventions and compare student performance in early foundation STEM courses across the same time frame.

Results: The binary classification model identified High-Risk students with an accuracy of 63–70% and an MCC of +0.28 to +0.30. In addition, we found statistically significant improvement ($p < 0.001$) in the STEM retention rates, 6-year graduation rates, and first part of Physics, Calculus, and Chemistry sequences after the interventions.

Discussion: The methodology and strategies presented may provide effective guidance for institutions seeking to improve the overall performance of undergraduate students who otherwise might struggle in their first-year engineering curriculum.

KEYWORDS

STEM, engineering curriculum, first year, persistence, retention, spatial visualization

OPEN ACCESS

EDITED BY
Hon-Ming Lam,
The Chinese University of Hong Kong, China

REVIEWED BY
Lau Victor,
The Chinese University of Hong Kong, China
Özcan Özyurt,
Karadeniz Technical University, Turkey

*CORRESPONDENCE
David P. Wick
dpwmcs@rit.edu
Emmanuel Atindama
atindaea@clarkson.edu
Prashant Athavale
prashant@clarkson.edu

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to STEM Education, a section of the journal Frontiers in Education

RECEIVED 01 November 2022
ACCEPTED 25 November 2022
PUBLISHED 09 December 2022

CITATION
Wick DP, Atindama E, Athavale P, Mondal S, Ramsdell MW, Jaspersohn RP and Moosbrugger JC (2022) Measuring the impact of student success retention initiatives for engineering students at a private research university. *Front. Educ.* 7:1087064. doi: 10.3389/feduc.2022.1087064

1. Introduction

Recognition of the importance of Science, Technology, Engineering, and Mathematics (STEM) as a national concern has been the subject of multiple reports over many years (Zumeta and Raveling, 2001; National Science Board, 2007; National Academy of Sciences, 2010; National Science Foundation, 2010a,b; Provasnik et al., 2012; National Research Council, 2013). At the highest level, the President's Council of Advisors on Science and Technology called for higher education institutions to produce more STEM graduates in order for the United States to remain competitive in the global economy (Olson and Riordan, 2012). Considering that approximately half of bachelor's degree-seeking students who enter STEM fields switch out of or fail to complete a STEM degree, increasing the retention of students who have already entered the STEM education pipeline is paramount to achieving this goal (Chen, 2013).

The reasons for students departing from STEM majors are varied, and intersecting factors make analyses as complex as the process of designing effective interventions. For example, STEM attrition rates are greater for women, historically underrepresented students of color, and first-generation students (Anderson and Kim, 2006; Griffith, 2010; Hill et al., 2010; Shaw and Barbuti, 2010). Non-cognitive factors such as motivation and self-efficacy also impact student success in STEM (Burtner, 2005; Al-Sheeb et al., 2019). Additionally, students who are less prepared academically for the challenge of a STEM curriculum have higher attrition rates (Astin and Astin, 1992; Shaw and Barbuti, 2010; Whalen and Shelley, 2010).

In this paper, we describe a methodology to identify and help under-prepared students entering undergraduate majors in STEM at a small technologically-oriented research university. In Section 2.1, we describe a binary classifier developed using student performance data over 5 years (2006 to 2010) to identify students who may be at a higher risk of failing in one or more early foundation STEM courses. We then used the classifier to advise interventions for engineering students categorized as "high risk" from 2011 to 2015. The primary intervention directed these students along an alternative curriculum pathway, intentionally decoupling the timing of first-year Calculus and Physics courses, traditionally taken simultaneously during the first semester. Additional interventions included mandatory enrollment of students in a Co-Calculus support course taken in tandem with Calculus and an optional training course offered to students identified with low spatial visualization skills. We investigate the following questions in this paper:

1. Is the binary classification methodology valid in identifying high-risk students?
2. Did the interventions have a positive impact on the graduation rates and STEM retention rates for high-risk engineering students?
3. Did the interventions have a positive impact on student success in the early foundation STEM courses?

1.1. First-year curriculum

The university uses a "Common-Core Curriculum" of first-year courses that provides engineering students with relatively flexible options for exploring and changing majors within the engineering disciplines. This work only considers first-year students without advanced placement credits for Calculus, Physics, or Chemistry. For such students, the traditional pathway through any discipline in engineering includes the following early foundation STEM courses taken during the first year: two sequential semesters each of Calculus (Calculus I and II), Calculus-based Physics (Physics I and II), and Chemistry (Chemistry I and II). Calculus is a co-requisite for Physics, meaning students need to take it with or prior to taking Physics. Physics and Chemistry have laboratory components and are, consequently, 4-credit courses. Calculus is a 3-credit course; however, Calculus I includes a complementary 2-credit Co-Calculus support course.

1.2. Challenges faced by under-prepared students in the first year

The rigid sequencing of required STEM courses in engineering can present challenges for students trying to navigate myriad pre and co-requisites successfully. The 'common curriculum' provides exploratory opportunities across engineering disciplines for most students who complete the core courses successfully and sequentially. However, the standard sequence can present obstacles that inhibit on-time progression for a portion of students who struggle during the first semester. For example, students who fail Physics I in the first semester typically retake it in the second semester, and if successful during their second attempt, still find themselves one course behind in their program. Such students have to face an overloaded subsequent semester or take a summer make-up course to avoid extending their time to degree completion. Even more detrimental is the case for students who fail Calculus I in the first semester. Even if they pass their Calculus-based Physics I course, they must retake Calculus I in the second semester and are prohibited from progressing to Physics II, which requires Calculus II as a co-requisite. Such students find themselves even further behind after just starting their engineering major. Furthermore, students who pass but score below a C grade are strongly encouraged by the engineering departments to retake Calculus and Physics as these courses are prerequisites for future courses in engineering and set the foundation for success in those courses. For reference, from 2006 to 2010, the percentage of all STEM students failing to achieve a C grade or better in their first semester of Physics and Calculus was approximately 25 and 33%, respectively (Jaspersohn, 2017).

1.3. Interventions recommended by the first-year council

In 2010, the university formed a First-Year Council to implement and coordinate strategies to improve student performance and retention in STEM majors. The council initiated a plan to identify each incoming student's preparedness for the first-year STEM experience by formalizing the collection of the following pre-college or pre-entry survey data. We list these instruments with generic labels for consistency in this article, with their formal titles and references in parentheses.

1. Math Diagnostic Survey (Clarkson University Math Skills Assessment) (Turner, 2008),
2. Physics Diagnostic Survey (FCI-Force Concepts Inventory) (Hestenes et al., 1992), and
3. Spatial Visualization Survey (Rotations component of the Purdue Spatial Visualization assessment) (Guay and McDaniel, 1977).

The Physics Diagnostic Survey provides a measure of conceptual understanding of Newtonian Physics without mathematical calculations, while the Math Diagnostic Survey assesses basic Mathematics skills relevant to beginning STEM majors. Combined, they provide predictive evidence of student performance in the early foundation STEM courses. The council identified the most under-prepared or "high-risk" students by comparing incoming student data (beginning in 2011) with historical data collected previously (2006–2010, pre-intervention years) and leveraging existing analyses. It used this information to inform the development of strategies going forward. Pre-enrollment measures capable of identifying each student's risk level allowed for targeted placement recommendations based on individual needs (2011–present, intervention years). The council enacted the following strategies in an attempt to improve student achievement in introductory STEM courses and increase retention and graduation rates in STEM majors:

1. **Alternative pathway:** The council provided engineering students identified in the high-risk category from Fall 2011 onward with an alternate schedule (strategic placement recommendation). In this schedule, the council moved Physics I to the second semester (and this consequently moved Physics II to the third semester) and replaced it with a required engineering course titled Engineering and Society (Moosbrugger et al., 2012; Chapman et al., 2015). The rationale was that by decoupling Physics I and Calculus I from the same semester, the least prepared engineering students would have a better chance of improving their mathematics skills before taking Calculus-based Physics. Essentially, this change delayed Physics by design for the students in the high-risk category instead of necessity (through failure, as was the case historically for many of these students) without sacrificing time for degree completion. Note that this change required some engineering departments to consider additional or customized extensions of the alternative pathway for their second-year courses that require Physics II as a pre/co-requisite. However, with the encouragement of the First Year Council, departments saw value in accommodating students preemptively by design as a proactive measure that reinforced the university's commitment to student success.
2. **Co-calculus for all:** Before 2011, the university automatically scheduled only students who scored low on the pre-entry Math Diagnostic Survey for the Co-Calculus support course, a low-credit mathematics skills course that complements Calculus. Since only low-scoring students were enrolled, a negative stigma was associated with this approach. On the council's recommendations, beginning in 2011, the university placed all Calculus students in Co-Calculus, regardless of their pre-entry score. Students were given the choice of remaining throughout the semester (to receive credit) or testing out once they achieved a normalized score of 0.90 or higher on a subsequent competency test (given nearly weekly), essentially shifting the course perception from a "fail in" model to a more positive "pass out" model. From 2011 to 2015, approximately 19% of the 3961 students enrolled in the first semester of Co-Calculus opted to forego credit and exit the course after achieving successful scores, including 24% of the 827 engineering students categorized as high-risk.
3. **Optional spatial visualization (sv) training:** Students whose normalized score fell below a selected cut-off (typically between 0.60 and 0.70) on the Rotations component of the Purdue Spatial Visualization assessment prior to entry were automatically scheduled for a one-semester SV training course (meeting once per week) from Fall 2012 onward. On the first training day, the students were encouraged to participate but were not required to remain. Each year, the cut-off varied to accommodate the reality of scheduling constraints (section capacity). From 2012 to 2015, approximately 13.6% of 1,851 incoming engineering students participated in Spatial Visualization training, including 19% of the 650 engineering students categorized as high-risk.

2. Materials and methods

2.1. Binary classification of student risk levels

Building upon historical data collected over multiple years (Turner, 2008), the First-Year Council initiated a comprehensive assessment of first-year performance in the

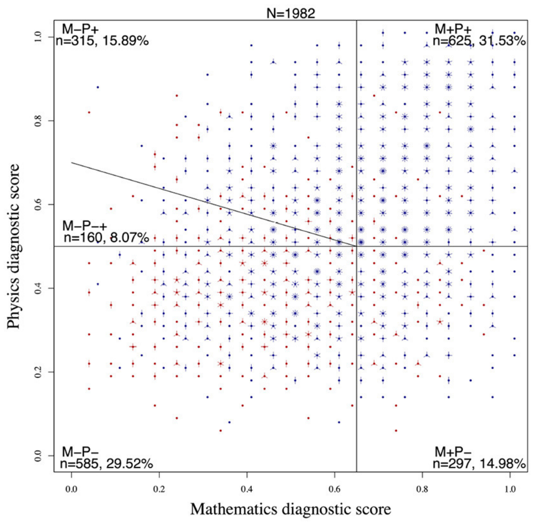

FIGURE 1
Sunflower plot of initial state data (Mathematics and Physics diagnostic scores) from 2006 to 2010 with group labels, counts, and percent of the total number of students. There were $N =$ 1,982 total students considered during this period who took the Mathematics and Physics diagnostic tests. The number n (lowercase) in each subgroup (M+P+, M−P+, M−P−+, M−P−, and M+P−) indicates the number of students in that subgroup. We also indicate the percentage of these numbers out of the total number N. Whenever data points with "S" and "U" grades overlapped, we added a slight offset for clarity.

introductory Calculus and Physics courses, supported by a 2009 grant from Procter & Gamble (P&G) (Schalk et al., 2009, 2011). The results of a Principal Component Analysis (PCA) of historical pre-entry or initial state data identified the Mathematics and Physics Diagnostic scores as relatively independent measures capable of explaining a significant amount of variance in the data (Schalk et al., 2009). To illustrate this, we created a sunflower plot (Figure 1) of historical initial state data (Fall 2006–2010 cohorts) of paired normalized Mathematics and Physics Diagnostic scores (N =1,982) for incoming students co-enrolled in both Calculus I and Physics I during their first semester. A solitary dot in the sunflower plot represents a single data point, whereas multiple petals represent multiple points at the same coordinate location. The blue and red colors, respectively, distinguish between students who were ultimately successful and those who were unsuccessful in their first semester Physics I course (as defined in Section 2.2).

The scatter in the data highlights the diversity of incoming student preparation levels. The sunflower plot divides the data into four quadrants based on students' performance on the Mathematics and Physics Diagnostic Surveys. We defined "success" on the Physics Diagnostic Survey as scoring 50% or more and "success" on the Mathematics Diagnostic Survey as 65% or more. We detail the rationale for these cutoffs in Section 3.2. Thus, the sunflower plot provides four general groups or quadrants indicating relative preparedness levels in Mathematics and Physics.

Each group, labeled with an "MP" for Math/Physics and "+" or "−" designation denoting a relative strength or weakness, respectively, represents a preparedness level. Thus, we have four risk categories: M+P+, M+P−, M−P+, and M−P−. Students considered to be well-prepared in both Mathematics and Physics are categorized as Low-Risk (M+P+), while students who are ill-prepared for both are considered High-Risk (M−P−). We categorized the students who are well-prepared in one but ill-prepared in the other as Medium-Risk (M+P− and M−P+). As a logistical control mechanism for maximizing enrollment in the Engineering and Society course during the intervention years, a small number of M− students who were just above the Physics cutoff were included in the Alternative Pathway recommendation for the Fall 2012 cohort and beyond. This additional group represents the highest risk students in the M−P+ Medium-Risk category and is labeled as M−P−+, shown as a wedge in Figure 1. Students in this subgroup were relabeled as high-risk, thus expanding the total count of students in the High-Risk category. Since the classification aims to identify high-risk students, we combine the low and medium-risk students into a single category, leading to binary classification.

2.2. Evaluation criteria for the classification model

Before using the classification model described in Section 2.1 in practice, we needed to validate the model's predictive capability. To this effect, we considered the final grades in the first-year foundational STEM courses relative to the identified risk categories for Fall 2006–2010 cohorts. We used the Receiver Operating Characteristics (ROC) curve to display the paired False Positive and True Positive Rates for students in Physics I, obtained by varying each cutoff or threshold between 0 and 1 by increments of 0.01. We label students earning a course grade of "C" or better in their first attempt as successful (S), while students earning below a "C" (including withdrawals, late withdrawals, and incompletes) in their first attempt as unsuccessful (U). In Figure 1, successful students appear blue, while unsuccessful students appear red. Since we are interested in identifying the High-Risk students with the classifier, the "Positive" instance is associated with identifying an unsuccessful student. Consequently, classifying a student into Low or Medium-Risk is labeled as a "Negative" instance. In the context of this binary classification, we define the True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) as follows:

True Positives (TP) = Students classified as high-risk who received "U" grade

False Positives (FP) = Students classified as high-risk who received "S" grade

True Negatives (TN) = Students classified as low and medium risk who received "S" grade

False Negatives (FN) = Students classified as low and medium risk who received "U" grade.

Consequently, we present the student counts as a set of 2 × 2 confusion matrices, where a "positive" instance is associated with identifying an unsuccessful student. Each confusion matrix includes the following model evaluation metrics (Swets, 1988; Fawcett, 2006) defined below:

$$\text{True Positive Rate (TPR)} = \frac{\text{Number of True Positives (\#TP)}}{\text{Number of students with "U" grade}}$$

$$\text{True Negative Rate (TNR)} = \frac{\text{Number of True Negatives (\#TN)}}{\text{Number of students with "S" grade}}$$

$$\text{False Positive Rate (FPR)} = \frac{\text{Number of False Positives (\#FP)}}{\text{Number of students with "S" grade}}$$

$$\text{False Negative Rate (FNR)} = \frac{\text{Number of False Negatives (\#FN)}}{\text{Number of students with "U" grade}}$$

$$\text{Positive Predictive Value (PPV)} = \frac{\text{Number of True Positives (\#TP)}}{\text{Number of students classified as high-risk students}}$$

$$\text{Negative Predictive Value (NPV)} = \frac{\text{Number of True Negatives (\#TN)}}{\text{Number of students classified as low and medium risk}}$$

$$\text{Accuracy (ACC)} = \frac{\text{Number of True Positives (\#TP)} + \text{Number of True Negatives (\#TN)}}{\text{Total number of students}}$$

$$\text{Matthews Correlation Coefficient (MCC)} = \frac{(\#TP)\cdot(\#TN) - (\#FP)\cdot(\#FN)}{\sqrt{(\#TP + \#FP)\cdot(\#TP + \#FN) + (\#TN + \#FP)\cdot(\#TN + \#FN)}}.$$

Accuracy (ACC) measures how well the classifier correctly identifies the categories. Accuracy varies from 0 to 1, with 1 indicating exact classification. Since we typically have fewer students in the High-Risk category than in the Low and Medium-Risk categories combined, we also calculated the Matthews Correlation Coefficient (MCC) (Matthews, 1975). The MCC, on a scale of −1 to 1, provides a measure of the overall quality of a binary prediction classifier. Positive values of the MCC indicate better prediction quality.

2.3. Comparability of high-risk students before and after intervention

Once the classification methodology was verified, as outlined in Section 2.2, the First-Year Council applied it to subsequent cohorts to make targeted placement recommendations. In Sections 2.4 and 2.5, we quantify the impact of the intervention on high-risk students from cohorts Fall 2011–2015 compared to the high-risk students from cohorts Fall 2006–2010. Since these cohorts are from different years, we needed to ensure they are indeed comparable. For a fair comparison, we needed the two groups to be similar with respect to their academic performance at the beginning of the first year at the university. We chose to use the SAT scores as a measure of the similarity between the two groups. We analyzed the descriptive statistics for math and verbal SAT scores of the two groups. Furthermore, we used the two-tailed *t*-test (Neter et al., 1996) on the SAT math and verbal scores of the two groups to compare them.

2.4. Analysis of the retention and graduation rates of engineering students

In this section, we narrow our analysis to just the High-Risk engineering-major students since the Alternative Pathway intervention was designed specifically for these students. We evaluated the long-term impact of the methodology and interventions implemented for engineering students identified as high-risk for Fall 2011–2015 cohorts.

2.4.1. Control and treatment groups for retention and graduation rates

For the analyses of the retention and graduation rates, we define the "control group" and "treatment group" as follows:

The control group is the group of engineering students identified as High-Risk students by the classifier described in Section 2.1, from Fall cohorts 2006 to 2010.

The treatment group is the group of engineering students identified as High-Risk students by the classifier described in Section 2.1 from Fall cohorts 2011 to 2015 who have received the treatment. To ensure that these students are only those who received the "treatment," we considered only the students enrolled in Physics I in the second semester of their cohort year and registered for the Engineering and Society course in the first semester of their cohort year.

We define first-year STEM retention as the percentage of first-year STEM major students enrolled in STEM majors at the beginning of their second year. We define second-year STEM retention similarly, as the percentage of first-year STEM major students enrolled in STEM majors at the beginning of their third year. We compared the first and second-year retention rates and the 4- and 6-year graduation rates before and after the interventions. All engineering programs at the university are 4-year programs. We also note that the 4-year graduation rate indicates an important "on-time" graduation rate metric, while the 6-year rate accounts for graduation within 150% of a student's program length. We used the two-population proportion test (Neter et al., 1996) to find the statistical significance of the difference between the two groups.

2.5. Measuring the impact of the interventions on early STEM courses

We wanted to quantify the effect of the interventions on success rates in the early STEM foundation courses taken by engineering students. Recall that we defined "success" as achieving a "C" grade or better on the first attempt. We compared the early STEM course performance of the students from the control group to the course performance of the students in the treatment group. We used the two-population proportion test to determine the statistical significance of the difference between the two groups.

2.5.1. Control and treatment groups for early STEM course success

For the control group in Physics I, Calculus I, and Chemistry I, we considered High-Risk engineering students in the cohort years 2006–2010 who took the corresponding courses in the first semester of their cohort years. Whereas, for the subsequent courses, Physics II, Calculus II, and Chemistry II, the 'control group' included High-Risk engineering students in the cohort years 2006–2010, who registered for these courses after passing the first part of the corresponding course. We only considered their Success or Failure in their "first" attempt at these courses for this study. The treatment group in Physics I, Calculus I, and Chemistry I consists of High-Risk engineering students in the cohort years 2011–2015 who registered for the Engineering and Society course in the first semester and Physics I course in the second semester of their cohort year. An additional requirement for students in treatment groups for Calculus I and Chemistry I is that the students in these groups need to have registered for these courses in the first semester of their cohort year. Whereas, for the subsequent courses, Physics II, Calculus II, and Chemistry II, the "treatment group" included High-Risk engineering students in the cohort years 2011–2015 who registered for these courses after passing the first part of the corresponding course.

2.6. Coding language and libraries used

We used Version 4.0.0 of R programming language for the coding with the following R libraries: readxl, dplyr, tidyr, ggplot2, and ggpubr.

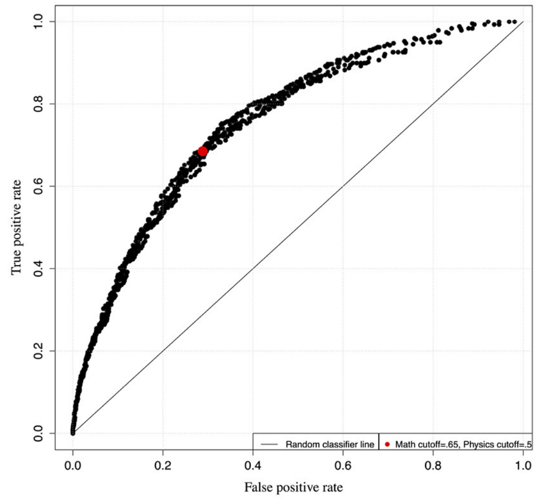

FIGURE 2
Receiver operating characteristics (ROC) curve for Satisfactory (S)/Unsatisfactory (U) grades in Physics I by predicted risk categories for Fall 2006–2010 cohorts.

3. Results

3.1. The binary classification model

In Figure 1 we depict the sunflower plot for all students from the Fall cohorts from 2006 to 2010. This figure is for the Mathematics Diagnostic survey cut-off of 0.65 and the Physics Diagnostic survey cut-off of 0.5. The binary classification system identified 745 students as high-risk students out of 1982. These 745 students include 585 students from the M−P− category and 160 students from the M−P−+ category as explained in Section 2.1.

3.2. Evaluation of the classification model

We depict the ROC curve for Physics I in Figure 2. Each dot in this figure plots the False Positive Rate (FPR) and True Positive Rate (TPR), corresponding to a pair of potential cutoff values in the Mathematics and Physics diagnostic surveys. This methodology implies that cutoffs may need to be decided for every course and for every cohort group. However, the primary intervention for engineering students involved shifting the timing of the first physics course by one semester. Hence, for practical and logistical reasons, the final cutoffs of 0.65 for the Mathematics and 0.50 for the Physics Diagnostic surveys were based on the ROC curve associated with Physics I, the impact of the expanded high-risk category (M−P−+), and guided by recommendations from the Mathematics and Physics Departments. The red dot in Figure 2 shows the corresponding

TABLE 1 Confusion matrices with student-counts demonstrating early STEM foundation course performance measured as Satisfactory (S)/Unsatisfactory (U) grade, by predicted risk categories for Fall 2006–2010 cohorts for the Math Diagnostic cut-off of 0.65 and Physics Diagnostic cut-off of 0.5.

		STEM performance by risk categories									
		Predicted risk level		Model evaluation metrics							
Actual grades		High	Medium + Low	TPR	TNR	FPR	FNR	PPV	NPV	ACC	MCC
Physics I Grade	U	285	(106 + 25)= 131	0.685	0.706	0.294	0.315	0.383	0.894	0.702	+0.329
	S	460	(506 + 600) = 1106								
Calculus I Grade	U	340	(126 + 29)= 155	0.689	0.638	0.362	0.313	0.427	0.838	0.625	+0.294
	S	456	(436 + 368) = 804								
Chemistry I Grade	U	283	(102 + 38)= 140	0.669	0.659	0.341	0.331	0.361	0.874	0.661	+0.277
	S	501	(482 + 486) = 968								

Success in coursework was defined as a letter grade of C (grade point 2.0) and above on the first attempt.

pair of FPR and TPR for these values and its proximity between the coordinate location (0,1), representing a perfect classifier, and the diagonal line representing a completely random classifier.

Table 1 summarizes the confusion matrices for the courses Physics I, Calculus I, and Chemistry I based on the final cutoff values. The proportion of students receiving a U grade correctly categorized as High-Risk (True Positive Rates) ranges from 0.67 to 0.69 for all three early STEM foundation courses, with False Positive Rates ranging from 0.29 to 0.36. The proportion of students receiving an S grade correctly categorized as Medium or Low-Risk (True Negative Rates) ranges from 0.64 to 0.71, with False Negative Rates ranging from 0.31 to 0.33. The proportion of students categorized as High-Risk who received a U grade ranges from 0.36 to 0.43, while the proportion of students categorized as Medium or Low Risk who received an S grade ranges from 0.84 to 0.89. The proportion of total students correctly categorized (Accuracy) ranges from 0.63 to 0.70. Since fewer students are typically in the high-risk category than in the low-medium risk category, we computed the Matthews correlation coefficient (MCC). We found that the MCC was positive and ranged from +0.28 to +0.33.

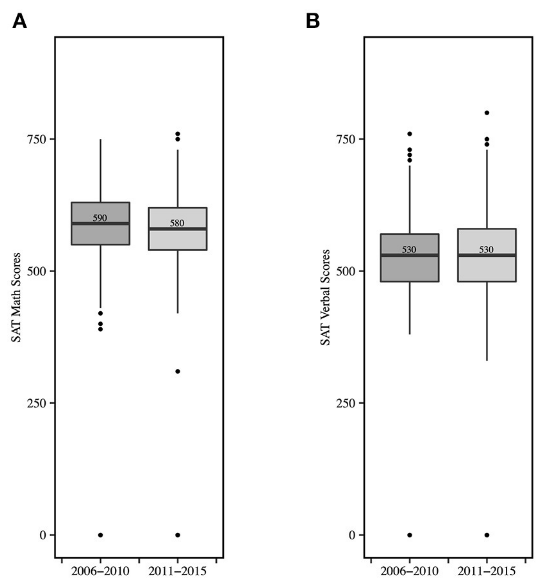

FIGURE 3
The comparison of the **(A)** SAT math scores, and **(B)** SAT verbal scores of the high-risk students from Fall 2006–2010 and the high-risk students from Fall 2010–2015. The *t*-tests yielded *p*-value > 0.05 in both cases, indicating no statistically significant difference between the two groups.

3.3. Results of the comparability of the high-risk students before and after intervention

We used SAT scores to verify the comparability of the high-risk students before and after the intervention. The mean SAT math scores of high-risk students from Fall 2006 to 2010 was 559, as opposed to 546 for the high-risk students from Fall 2011 to 2015. The first quartile, median, and third quartiles for SAT math scores were 550, 590, and 630, respectively, for the high-risk students from Fall 2006 to 2010, whereas these statistics were 540, 580, and 620, respectively, for the high-risk students from Fall 2011 to 2015. The *t*-test showed that the SAT math scores of the two groups were statistically similar ($p = 0.126$). The mean SAT verbal score of high-risk students from Fall 2006 to 2010 was 506, compared to 503 for the high-risk students from Fall to 2015. The first quartile, median, and third quartiles for SAT verbal scores were 480, 530, and 570, respectively, for the high-risk students from Fall 2006 to 2010, whereas these statistics were 480, 530, and 580, respectively, for the high-risk students from Fall 2011 to 2015. We summarize the descriptive statistics in Figure 3. Furthermore, the *t*-test revealed that the SAT verbal scores of the two groups were statistically similar

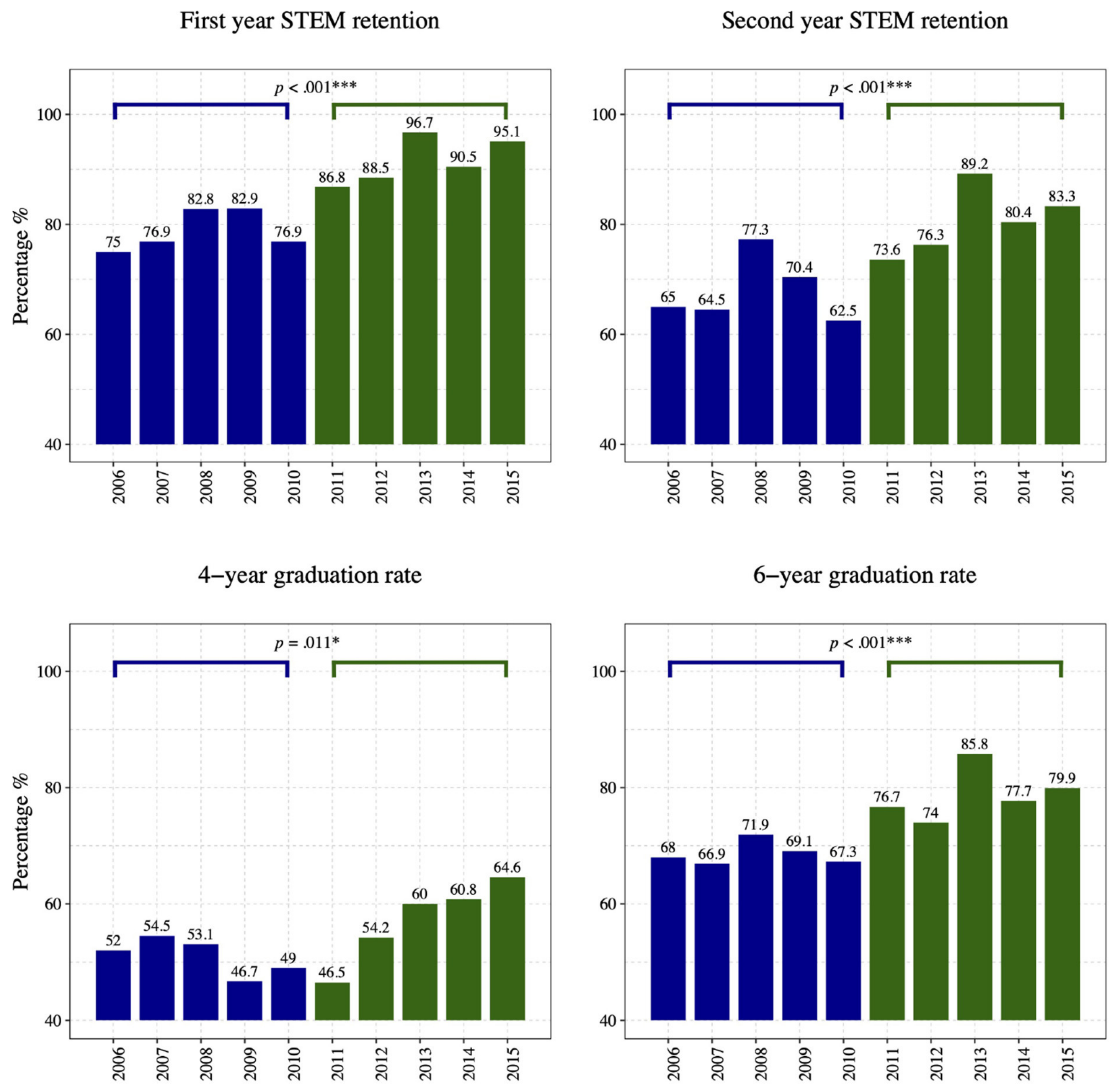

FIGURE 4
The first row shows the High-Risk engineering students' first and second-year retention rates in the STEM major for the cohort years Fall 2006–2015. The second row depicts the 4- and 6-year graduation rates. The percentage values for each year are indicated at the top of each bar in the bar charts. The results from the pre-intervention years (Fall 2006–2010) are shown in blue, whereas those from post-intervention years (Fall 2011–2015) are shown in green. The significance levels (p-values) indicating the differences in the pre-intervention and post-intervention years are shown in the individual graphs. ***$p < 0.001$, **$p < 0.01$, and *$p < 0.05$.

($p = 0.706$). Since, in both cases, the $p > 0.05$, we could compare the performance of these two groups.

3.4. Retention and graduation rate comparison

Figure 4 shows the first and second-year retention and the 4- and 6-year graduation rates for engineering students identified as High-Risk for each Fall entry cohort from 2006 to 2015. We can see that the High-Risk students' overall retention and graduation rates are better during the intervention years than before the interventions. The plots also show the p-values for the two-population proportion tests comparing the control and treatment groups. We see that the treatment group performed significantly better than the control group in the first and second-year retention rates with $p < 0.001$. The 4- and 6-year graduation rates in the treatment group were also significantly better than the control group with $p < 0.05$ and $p < 0.001$, respectively. Table 2 shows the details of these tests, including the overall percentage of students in each group.

3.5. Comparison of success rates in early STEM courses

Figure 5 depicts the success rates in early STEM courses for engineering students identified as High-Risk for each Fall cohort from 2006 to 2015. From the bar chart for Physics I, we can readily recognize that the treatment group performed better than the control group. In fact, the minimum success rate in the post-intervention years (71.8%) is significantly greater than the maximum success rate in the pre-intervention years (66.3%). However, we can not reach the same conclusion for other courses without further investigation. Thus, we performed the two-population proportion test to compare the two groups' performances for each early STEM foundation course. We display the p-values for the two-population proportion tests to compare the control and treatment groups in Figure 5. Table 3 shows the details of these tests, including the overall percentages and number of students in each group. Note that we used the number of students attempting the courses for the first time in these tests. The treatment group performed significantly better ($p < 0.001$) in Physics I, Calculus I, and Chemistry I than the control group. For Calculus II, we found a marginal improvement ($p < 0.05$) in the success rate of the treatment group over the control group. However, we found no statistically significant difference between the two groups' performance in Physics II and Chemistry II.

4. Discussion

We used the historical data from the Fall cohorts from 2006 to 2010 to develop and refine a model classifier to identify students at high risk of underperforming in the early STEM courses. While the classification method has limitations and room for improvement (see Section 4.1), as a low-dimensional model based on only two relatively independent measures easily captured at the pre-entry point of enrollment, the results are promising. The relatively high accuracy for correct categorization seems in contrast to the somewhat low Matthews Correlation Coefficient for overall model quality. However, we find much value in using this approach for identifying the majority of genuinely high-risk students. The high Negative Predictive Values suggest that in using this methodology as a predictive tool, we should have high confidence that most students we categorize as Medium/Low-Risk will likely do well in the early foundation STEM courses. Furthermore, while the Positive Predictive Values are much lower due to the misclassification of a fair number of successful students as

TABLE 2 Success rates for the control group and the treatment group along with the p-values for the two proportion tests.

	Control group ($N_1 = 605$)	**Treatment group ($N_2 = 672$)**	**p-values**
First-year STEM retention	86.1%	93.6%	<0.001***
Second year STEM retention	76.4%	85.4%	<0.001***
4-year graduation rate	50.9%	57.4%	0.011*
6-year graduation rate	68.8%	78.7%	<0.001***

There were $N_1 = 605$ students in the control group. Moreover, there were $N_2 = 672$ students in the treatment group. See Section 2.4.1 for the definitions of the control and treatment groups. $^{***}p < 0.001$, $^{**}p < 0.01$, and $^{*}p < 0.05$.

High-Risk, the model correctly classifies more than 2/3 of the unsuccessful students (in Physics I, Calculus I, and Chemistry I) as High-Risk. From an intervention design perspective, this aligns with a conservative approach in which we may offer more students additional assistance or recommendations to enhance their success than may be necessary.

After verification, we used the classifier to identify High-Risk engineering students in the Fall cohorts of 2011 onward. These students were then prescribed an alternative pathway, decoupling the concurrent timing of the Calculus I and Physics I courses. Additionally, all students were enrolled in a Co-Calculus support course, and some were provided with an optional SV training course. We found that these combined interventions had a statistically significant ($p < 0.01$) positive impact on the 4- and 6-year graduation rates and the first and second-year STEM retention rates of the High-Risk students. The interventions also improved the performance of the High-Risk students ($p < 0.001$) in the courses Physics I, Calculus I, and Chemistry I. The improvement in the Physics I success rate could be attributed to these students completing the Calculus I course before attempting Physics I. This finding contrasts with the pre-intervention years when all students took Calculus I and Physics I concurrently. We note that improvements in the success rate in Calculus II were marginally significant and not significant for Physics II and Chemistry II. However, this observation may reflect the impact of increased retention of more High-Risk students in the post-intervention years.

4.1. Limitations and future work

The approach presented is understandably limited in that it does not directly account for non-cognitive factors and demographic variables linked to overall student success. However, it uses relatively easy-to-capture diagnostic data at the pre-entry point of student enrollment. The resulting binary classifier is admittedly "static" in that the cut-offs for the Mathematics/Physics Diagnostic surveys are predetermined by

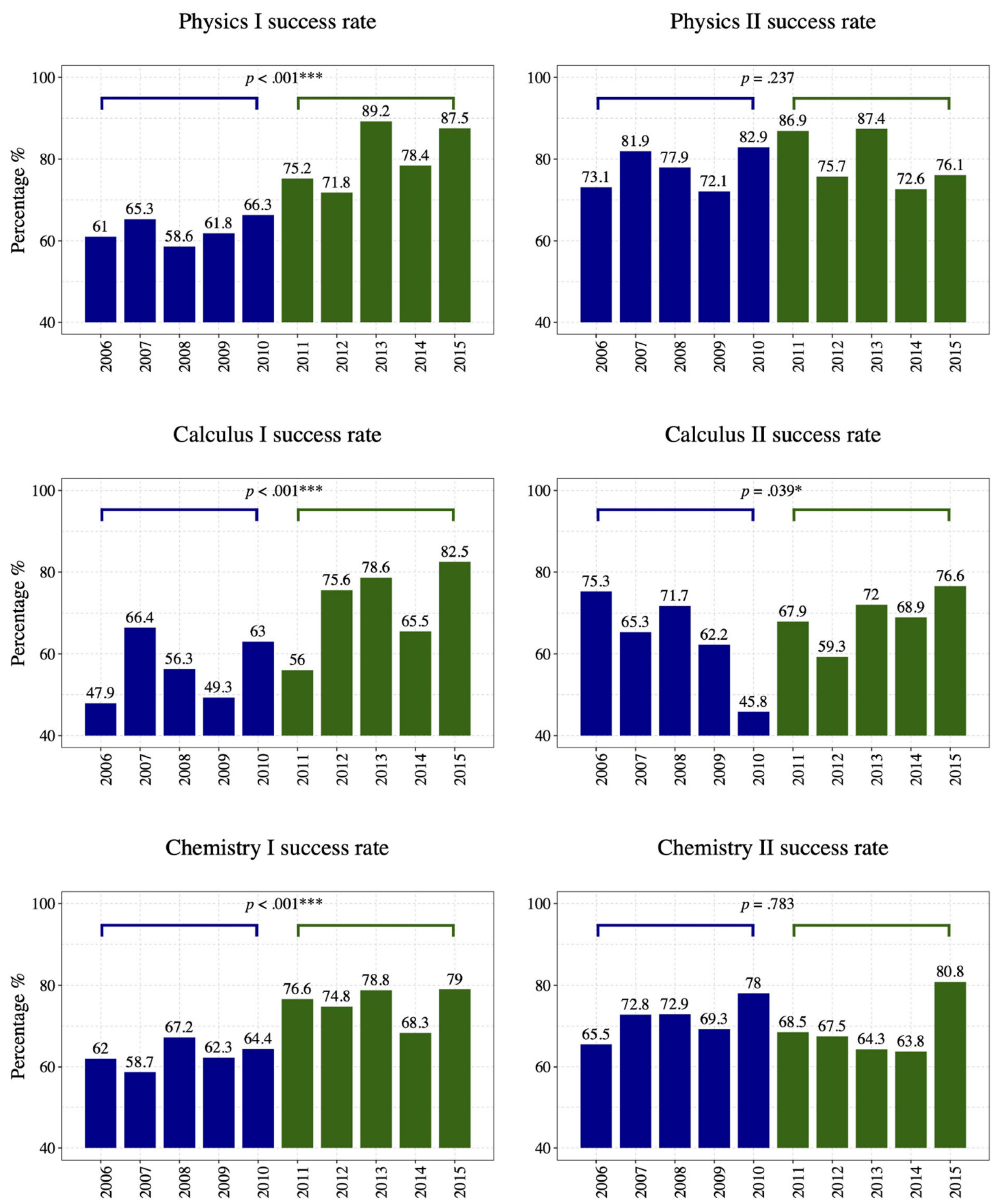

FIGURE 5
The success rates of High-Risk engineering students in the early STEM courses from the cohort years Fall 2006–2015. The success rates in the pre-intervention years (Fall 2006–2010) are shown in blue, whereas the post-intervention years (Fall 2011–2015) are in green. The significance levels and the *p*-values indicating the differences in the pre-intervention and intervention years are shown in the individual graphs. $^{***}p < 0.001$, $^{**}p < 0.01$, and $^{*}p < 0.05$.

TABLE 3 Success rates for the control group and the treatment group (as a percentage and raw numbers) along with the *p*-values for the two proportion tests.

	Control group	Treatment group	*p*-values
Success in Physics I	62.5% (378)	80.4% (540)	$<0.001^{***}$
Success in Calculus I	56.4% (332)	71.7% (466)	$<0.001^{***}$
Success in Chemistry I	62.9% (378)	75.3% (495)	$<0.001^{***}$
Success in Physics II	77.3% (371)	79.3% (456)	0.237
Success in Calculus II	64.0% (310)	69.2% (430)	0.039^{*}
Success in Chemistry II	71.5% (352)	69.2% (413)	0.783

Success in coursework was defined as a letter grade of C (grade point 2.0) and above on the first attempt. See Section 2.5.1 for the definitions of the control and treatment groups. $^{***}p < 0.001$, $^{**}p < 0.01$, and $^{*}p < 0.05$.

the data from 2006 to 2010. In the short term, i.e., for the period 2011 to 2015 considered in the paper, this model worked reasonably accurately and as expected. However, in the future, we will need to reexamine the long-term validity of the model and make adjustments accordingly.

Moreover, students whose scores are near the intersection of the Mathematics and Physics Diagnostic cut-offs shown in the sunflower plot are similar in preparedness but could essentially belong in any of the four categories. A revised set of cut-offs with diagonal (negative slope) or curved diagonal bands that broaden the Medium-Risk categories into a single zone while simultaneously separating the High and Low-Risk zones might improve the predictive capability and enhance targeted recommendations for further improving retention and graduation rates. Additionally, a future study incorporating modeling techniques will examine the extent to which each intervention strategy contributed to student success.

5. Conclusion

In this paper, we designed a binary classifier to identify students at higher risk of underperforming in early foundation STEM courses. We used student performance data from 2006 to 2010 to design the classification model and validated it using a classifier accuracy measure and Matthews Correlation Coefficient. After the validation, we used this model to identify the most underprepared engineering students from subsequent incoming cohorts. Once identified, these students were prescribed interventions (alternative pathways, a Co-Calculus support course, and an optional SV training course) to help them succeed in their engineering programs. We observed that these collective interventions significantly and positively impacted the STEM retention rates of these students in the first 2 years of their academic careers and improved their 4- and 6-year graduation rates. Moreover, the performance of the High-Risk engineering students also improved in the early foundation STEM courses, translating to increased retention.

These findings provide an effective methodology for identifying and supporting engineering students likely to struggle in their undergraduate education. Institutional profiles and student preparedness levels can vary significantly from one university to another. Hence the methodology and suggested interventions may not translate directly with the same level of effectiveness for other institutions. However, the overall improvement in the graduation, retention, and success rates achieved in the early STEM courses suggests that customized analysis and targeted interventions can elevate student success. The strategies presented in this article may provide effective guidance for institutions seeking to improve the overall performance of undergraduate students who otherwise might struggle in their engineering curriculum.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Institutional Review Board, Clarkson University. Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

DW: writing, review, and supervision. EA: coding, data curation, investigation, analysis, and visualization. PA: writing, editing, data curation, and statistical analysis. SM: validation and review. MR: data curation and review. RJ: coding and data organization. JM: review. All authors contributed to the article and approved the submitted version.

Acknowledgments

We are grateful for the many years of support from Dr. Anthony G. Collins, the former President of Clarkson University. We acknowledge the significant contributions of Dr. Peter Turner, former Dean of the School of Arts & Sciences, Dr. Goodarz Ahmadi, interim Provost and former Dean of the

Coulter School of Engineering, and other faculty who served on the First-Year Council at Clarkson University.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Al-Sheeb, B. A., Hamouda, A., and Abdella, G. M. (2019). Modeling of student academic achievement in engineering education using cognitive and non-cognitive factors. *J. Appl. Res. Higher Educ.* 11, 178–198. doi: 10.1108/JARHE-10-2017-0120

Anderson, E., and Kim, D. (2006). *Increasing the Success of Minority Students in Science and Technology*. Washington, DC: American Council on Education.

Astin, A. W., and Astin, H. S. (1992). *Undergraduate Science Education: The Impact of Different College Environments on the Educational Pipeline in the Sciences. Final Report*. Higher Education Research Institute, University of California, Los Angeles, CA (ERIC Document Reproduction Service No. ED 362 404).

Burtner, J. (2005). The use of discriminant analysis to investigate the influence of non-cognitive factors on engineering school persistence. *J. Eng. Educ.* 94, 335–338. doi: 10.1002/j.2168-9830.2005.tb00858.x

Chapman, E. A., Wultsch, E. M., DeWaters, J., Moosbrugger, J. C., Turner, P. R., Ramsdell, M. W., et al. (2015). "Innovating engineering curriculum for first-year retention," in *2015 ASEE Annual Conference and Exposition* (Seattle, WA), 26–967.

Chen, X. (2013). STEM Attrition: *College Students' Paths into and out of STEM Fields*. Statistical Analysis Report. NCES 2014-001. National Center for Education Statistics.

Fawcett, T. (2006). An introduction to roc analysis. *Pattern Recogn. Lett.* 27, 861–874. doi: 10.1016/j.patrec.2005.10.010

Griffith, A. L. (2010). Persistence of women and minorities in stem field majors: is it the school that matters? *Econ. Educ. Rev.* 29, 911–922. doi: 10.1016/j.econedurev.2010.06.010

Guay, R. B., and McDaniel, E. D. (1977). The relationship between mathematics achievement and spatial abilities among elementary school children. *J. Res. Math. Educ.* 8, 211–215. doi: 10.5951/jresematheduc.8.3.0211

Hestenes, D., Wells, M., and Swackhamer, G. (1992). Force concept inventory. *Phys. Teach.* 30, 141–158. doi: 10.1119/1.2343497

Hill, C., Corbett, C., and St Rose, A. (2010). *Why so Few? Women in Science, Technology, Engineering, and Mathematics*. Washington, DC: AAUW.

Jaspersohn, R. P. (2017). *Addressing Student Success and Retention in STEM Majors Through Strategic Curriculum Pathways and Early Research Experiences*. Potsdam, NY: Clarkson University.

Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of t4 phage lysozyme. *Biochim. Biophys. Acta* 405, 442–451. doi: 10.1016/0005-2795(75)90109-9

Moosbrugger, J. C., and DeWaters, J., and Richards, M. C., and Chapman, E. A. (2012). "A course on engineering and society for first-year engineering students and non-majors," in *2012 ASEE Annual Conference and Exposition*, (San Antonio, TX). 25–34.

National Academy of Sciences (2010). *Rising Above the Gathering Storm, Revisited: Rapidly Approaching Category 5*. Washington, DC: The National Academies Press.

National Research Council (2013). *Monitoring Progress Toward Successful K-12 STEM Education: A Nation Advancing*? Washington, DC: National Academies Press.

National Science Board (2007). *A National Action Plan for Addressing the Critical Needs of the US Science, Technology, Engineering, and Mathematics Education System*. Alexandria, VA: National Science Foundation.

National Science Foundation (2010a). *Science and Engineering Indicators 2010. NSB 10-01*. Alexandria, VA: National Science Foundation.

National Science Foundation (2010b). *Preparing the Next Generation of STEM Innovators: IDENTIFYING and Developing our Nation's Human Capital*. Arlington, VA: National Science Foundation.

Neter, J., Kutner, M. H., Nachtsheim, C. J., and Wasserman, W. (1996). *Applied Linear Statistical Models, 4th Edn*. Chicago: Irwin.

Olson, S., and Riordan, D. G. (2012). *Engage to Excel: Producing One Million Additional College Graduates With Degrees in Science, Technology, Engineering, and Mathematics. Report to the President*. Washington, DC: Executive Office of the President.

Provasnik, S., Kastberg, D., Ferraro, D., Lemanski, N., Roey, S., and Jenkins, F. (2012). Highlights from TIMSS 2011: Mathematics and Science Achievement of US Fourth-and Eighth-Grade Students in an International Context. *NCES 2013–009*.

Schalk, P. D., Wick, D. P., Turner, P. R., and Ramsdell, M. W. (2009). "Impact: integrated mathematics and physics assessment for college transition," in *2009 39th IEEE Frontiers in Education Conference* (San Antonio, TX: IEEE), 1–6.

Schalk, P. D., Wick, D. P., Turner, P. R., and Ramsdell, M. W. (2011). "Predictive assessment of student performance for early strategic guidance," in *2011 Frontiers in Education Conference (FIE)*, (Rapid City, SD), S2H-1-S2H-5.

Shaw, E. J., and Barbuti, S. M. (2010). Patterns of persistence in intended college major with a focus on stem majors. *NACADA J.* 30, 19–34. doi: 10.12930/0271-9517-30.2.19

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science* 240, 1285–1293. doi: 10.1126/science.3287615

Turner, P. R. (2008). A predictor-corrector process with refinement for first-year calculus transition support. *Primus* 18, 370–393. doi: 10.1080/10511970601131639

Whalen, D. F., and Shelley, M. C. (2010). Academic success for stem and non-stem majors. *J. STEM Educ.* 11, 45–60. Available online at: https://www.jstem.org/jstem/index.php/JSTEM/article/download/1470/1292

Zumeta, W., and Raveling, J. (2001). *The Best and Brightest for Science*. Boston, MA: Springer U.S.

Frontiers in Education

TYPE Original Research
PUBLISHED 16 March 2023
DOI 10.3389/feduc.2023.1127984

Check for updates

OPEN ACCESS

EDITED BY
Wang-Kin Chiu,
The Hong Kong Polytechnic University,
China

REVIEWED BY
Vanessa Camilleri,
University of Malta,
Malta
Mawardi – Mawardi,
Padang State University,
Indonesia

*CORRESPONDENCE
Christine Lourrine S. Tablatin
✉ tablatinchristine@gmail.com

SPECIALTY SECTION
This article was submitted to
STEM Education,
a section of the journal
Frontiers in Education

RECEIVED 20 December 2022
ACCEPTED 27 February 2023
PUBLISHED 16 March 2023

CITATION
Tablatin CLS, Casano JDL and
Rodrigo MMT (2023) Using Minecraft to
cultivate student interest in STEM.
Front. Educ. 8:1127984.
doi: 10.3389/feduc.2023.1127984

Using Minecraft to cultivate student interest in STEM

Christine Lourrine S. Tablatin[1,2]*, Jonathan D. L. Casano[1] and Maria Mercedes T. Rodrigo[1]

[1]Ateneo Laboratory for the Learning Sciences, Department of Information Systems and Computer Science, Ateneo de Manila University, Quezon City, Philippines, [2]Information Technology Department, Pangasinan State University, Urdaneta City, Philippines

Due to the popularity and flexibility of Minecraft, educators have used this game to develop instructional materials and activities to cultivate student interests in science, technology, engineering and mathematics (STEM). One example of such an initiative is the What-If Hypothetical Implementations in Minecraft (WHIMC) project of the University of Illinois Urbana Champaign. The study reported in this paper describes a WHIMC deployment in the Philippines and the effects this deployment had on student STEM interest. The study used quantitative and qualitative methods to determine the effect of WHIMC on the STEM interest of Filipino students. We performed quantitative analysis of the pre- and post-STEM Interest Questionnaire (SIQ) ratings and Game Experience Questionnaire (GEQ) ratings of the high- and low-performers to determine the effect of using WHIMC in the students' STEM interest and the difference between the game experience of high- and low-performers, respectively. Qualitative analysis of the answers to the open-ended questions about the attributes of the module was also conducted to determine the relationship between the module attributes and student performance. The analysis of the aggregated SIQ ratings before and after using the WHIMC-based modules revealed only a minimal effect on the STEM interests of the students. However, there was a significant effect in the Choice Actions construct, which implies that students recognize the importance of studying hard if they want to pursue STEM-related careers. Further, the analysis of the overall GEQ of high-performers and low-performers also revealed no significant difference. Although no significant difference was observed in the overall GEQ, high-performers had significantly higher GEQ ratings in the Immersion dimension. This result suggested that high-performers had a more positive, engaging, and enjoyable learning experience. Moreover, the findings on the favorite module attributes suggested that students perform better in the out-of-game assessments when they like all the module attributes. This implies that students must be engaged in the game and learning task aside from being interested in the learning topic to have better assessment scores. The study also showed that open-ended learning environments coupled with tasks that demand exploration, observation, and higher-ordered thinking are demanding even on high-performers.

KEYWORDS

Minecraft, WHIMC world, STEM interest, digital game-based learning, educational games

1. Introduction

Learners often find STEM difficult because it requires complex thinking, repeated practice, and self-discipline (Bertozzi, 2014). According to the PISA National Report on the Philippines, compared to the OECD average of 489 in Math and 489 in Science, Filipino students scored a low 353 and 357, respectively. Only 1 out of 5 attained the minimum proficiency level in math (Education GPS, OECD, February 2023). These results are corroborated by students' performance in the National Achievement Test, where only 25% demonstrated mastery levels in math and only 5% of test takers demonstrated mastery levels in science. Thus, addressing STEM interest and achievement in the Philippines is an acute need. Improving students' self-efficacy through learning experiences is essential to cultivating students' interest and enthusiasm in STEM careers (Mohtar et al., 2019). One of the innovative ways to provide an engaging learning environment that keeps students interested and enthusiastic about STEM subjects is the use of games in learning.

Digital game-based learning (DGBL) has become a growing educational trend in the classroom as an engaging teaching approach for improving student motivation and learning (Ennis, 2018; Leong et al., 2018; Hussein et al., 2019; Shang et al., 2019). Games provide more amusement, enjoyment, and aesthetic appeal (Alawajee, 2021). They can also encourage the player to learn, offer multisensory environments, and improve the capacity of a player to think and create meaning (Iliya and Jabbar, 2015). The use of digital games can help students gain a more concrete understanding of abstract, theoretical topics while interacting with the learning material (Nkadimeng and Ankiewicz, 2022). Games have been advantageous for learning in different domains, including more authentic learning and increased student engagement because of their degree of interactivity and immersion (Alonso-Fernandez et al., 2019). Since STEM subjects are complex and challenging to learn, games can be a great way to introduce learners to scientific concepts. Several studies have demonstrated the beneficial impacts of games on science education. A study on DGBL for elementary science education revealed increased student engagement, domain knowledge, and problem-solving skills (Lester et al., 2014). Students who played the personalized DGBL application about plants gained a significant increase in learning achievements and motivation (Hwang et al., 2012). In addition, students who learned about migratory bird identification with the DGBL environment have significantly outperformed their peers in the acquisition of learning and motivation (Chu and Chang, 2014). The game Sorceress of Seasons was utilized to teach fundamental programming concepts. This resulted in increased positive attitudes toward programming, with female students reporting larger increases in computer science interest than males. The study suggests that games may be successful in increasing interest in STEM (Bonner and Dorneich, 2016). Further, the simultaneous presence of learning experiences and player self-determination while playing a STEM digital game might foster STEM interest (Ishak et al., 2022). The positive effects of digital games on student achievement, skills acquisition, motivation, and engagement have influenced educators, game developers, funding organizations, and researchers to use games across many platforms to teach STEM subjects (Bertozzi, 2014).

Minecraft is one of the game platforms used to teach and encourage interest in STEM. Minecraft is a sandbox-style video game released in 2009 by Mojang and the most widely played game in the world, with more than 180 million copies sold to date (Bitner, 2021). Due to its popularity and flexibility, educators utilize this game platform to develop instructional materials and activities to cultivate student interest in STEM. Pusey and Pusey (2015) used MinecraftEdu as an instructional tool to teach Earth Science topics to Grade 8 students in 2 schools in Australia. Along with the traditional teaching methods such as worksheets, slideshows, videos, and hands-on activities, the MinecraftEdu lessons were utilized once a week throughout the 5 to 6-week Earth Science program. Students who participated in the program expressed increased enthusiasm about attending science class because they liked the interactive learning, teamwork, and enjoyable coursework. This result showed that after the use of MinecraftEdu lessons, student interest in science has increased. Nkadimeng and Ankiewicz (2022) also reported a similar finding about using MinecraftEdu for a series of five 1 h lessons in atomic structure in a South-African junior high school. Further, learning with MinecraftEdu makes abstract concepts easier to understand, promotes critical thinking, and is conducive to collaboration and motivation. Another study prepared four different STEM activities and asked 6th-grade science classes to use Minecraft Educational Edition for 4 hours per week. The researchers collected data on STEM interests using the STEM Career Interest Survey and Scientific Creativity Scale. Both scientific creativity and STEM interest levels statistically increased (Saricam and Yildirim, 2021). These results imply that MinecraftEdu might be suitable as a learning tool for Science and Chemistry subjects. Furthermore, there is evidence from prior studies that games have a positive effect on STEM interests. However, there is a lack of longitudinal studies. Indeed, papers such as those of Gao et al. (2020) call for longitudinal studies to determine game-based learning's far-reaching effects.

What-If Hypothetical Implementations in Minecraft (WHIMC; https://whimcproject.web.illinois.edu/) also aims to engage, excite, and generate interest in learning science. WHIMC is a set of Minecraft worlds teachers can utilize as supplementary activities in teaching STEM. It includes a Rocket Launch Facility, the Lunar Base LeGuin, and a Space Station as shown in Figure 1. It also includes exoplanets and different versions of Earths, e.g., Earth with no moon, Earth with a colder sun. WHIMC immerses learners in simulated environments wherein they can move around these different worlds and make observations while exploring them (Yi and Lane, 2019; Manahan and Rodrigo, 2022; What-If Hypothetical Implementations in Minecraft (WHIMC), n.d.).

WHIMC has been the platform for several studies. One such study conducted during a summer camp examined campers' actions by giving them a quick 10 min presentation on hypothetical earth scenarios before allowing them to explore worlds in Minecraft. It revealed that sandbox games can spark interest in STEM subjects among underrepresented adolescents and that engagements with natural phenomena are possible in an open digital environment (Yi et al., 2020). Another study (Yi et al., 2021) examined interest triggers within Minecraft and found that personal relevance relates to a desire to reengage in camp content and with the design and structure of the intervention. Further study on STEM interest triggers within Minecraft in a hybrid summer camp found that various in-game and contextual aspects of the learning experiences, such as instructional conversation, novelty, ownership, and challenge, triggered the learners' STEM interests (Lane et al., 2022).

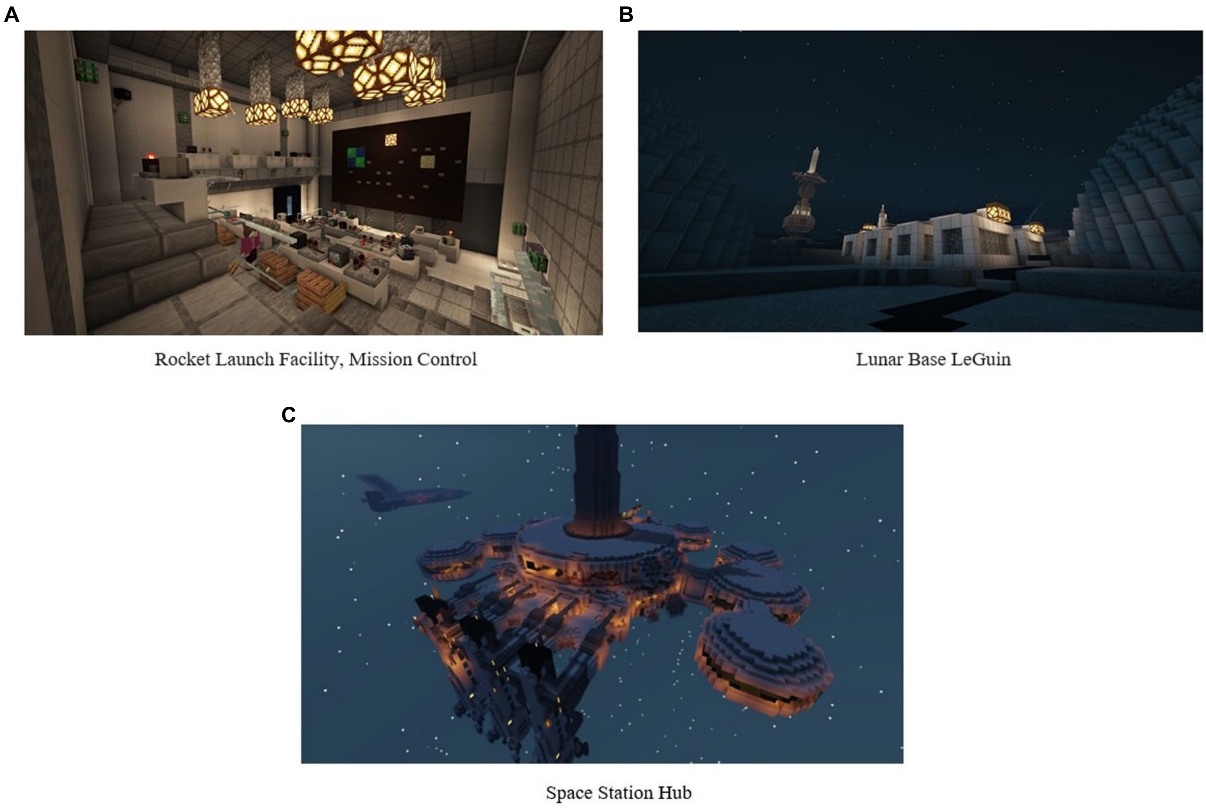

FIGURE 1
WHIMC worlds.

Gadbury and Lane (2022) encouraged teenagers to participate in five after-school sessions over the course of 5 weeks, during which they used Minecraft to explore several versions of Earth. The research investigates how different levels of STEM interest affect in-game science tool usage and observations across the hypothetical versions of Earth. The result revealed that participants with moderate STEM interests had the highest science tool usage, indicating high engagement and desire to learn. In terms of observations, participants with high STEM interests recorded high observations, suggesting confidence or high prior knowledge. Studies on the use of WHIMC were also conducted in the Philippines. The analysis of learner traversals of Minecraft worlds conducted in a grade school found a negative correlation between learner performance and overall distance traveled. This finding implied that low performers wander early in gameplay while high performers use a depth-first search strategy when exploring an area and are goal-oriented (Esclamado and Rodrigo, 2022a). The study of Casano and Rodrigo (2022a) performed a comparative assessment of American and Filipino learner traversals and in-game observations within Minecraft against canonical answers from experts. The finding suggested that high performers make more observations aligned with canonical answers from experts than low performers. They also found a difference in the in-game behavior of low performers. Filipino students tend not to make in-game observations, while American students actively make in-game observations. Another study looked at the achievement, behaviors, and STEM interests of frustrated and bored learners using WHIMC and found that frustrated learners tend to disengage from the game and bored learners tended to perform poorly on post-game assessments (Esclamado and Rodrigo, 2022b). Further, the analysis of game experience and STEM interest of primary school learners in the Philippines reported that high and low performers had the same level of game experience and that they like the game and learning-related WHIMC features. However, the learning task integrated into the WHIMC-based modules made learning difficult for the high performers, and technical bugs made learning difficult for the low performers. The finding on the STEM interest showed that high performers had a higher degree of agreement with the Stem Interest Questionnaire (SIQ) compared to the low performers (Casano and Rodrigo, 2022b). This study aims to continue the Philippine studies by promoting the use of WHIMC as a learning tool in a Philippine middle school to cultivate student STEM interests. Specifically, we seek answers to the following research questions:

> *RQ1*: What is the effect of using WHIMC on the STEM interests of students?
>
> *RQ2*: What is the difference between the game experience of high- and low-performers?
>
> *RQ3*: What is the relationship between the module attributes and student performance?

2. Materials and methods

The study used quantitative and qualitative methods to determine the effect of WHIMC on the STEM interests of Filipino students. We used an embedded design wherein we collected quantitative data from the SIQ and Game Experience Questionnaire (GEQ) survey questionnaires and qualitative data from the open-ended questions included in the GEQ questionnaire. Insights drawn from analyzing the answers to the open-ended questions about the module attributes might support the observations from the quantitative analysis of the SIQ and GEQ ratings. Therefore, we first performed quantitative analysis of the pre-SIQ and post- SIQ ratings and GEQ ratings of the high- and low-performers to determine the effect of using WHIMC in the students' STEM interests and the difference between the game experience of high- and low-performers, respectively. We then performed a qualitative analysis of the answers to the open-ended questions on the attributes of the module to determine the relationship between the module attributes and STEM interests. The research protocol was reviewed and approved by the University Research Ethics Committee of the Ateneo de Manila University.

2.1. Teacher-created learning modules

The research team established a formal partnership with the University of Illinois Urbana-Champaign (UIUC) team to gain access to WHIMC's content, code, and configuration details. A parallel server was then set up in the Philippines to run the experiments and manage the tasks without constantly coordinating with the UIUC team. After that, the research team established institutional partnerships with elementary and middle schools in the Philippines. Partner teachers were recruited, informed about the project goals, and requested to design WHIMC-based learning modules and out-of-game assessments. The research team gave the partner teachers 30 days to explore the WHIMC worlds to familiarize themselves with the game. The partner teachers then chose specific topics within their respective academic curriculum levels where they thought a particular WHIMC world would fit. The partner teachers and the research team reviewed the learning modules for quality, viability, and curriculum alignment before using these modules in class. The research team then provided documentation to assist the partner teachers in preparing for the WHIMC module implementation. The project manager also gave the partner teachers Minecraft account credentials to be used by the participating students in their class before the module implementation. Only the partner teacher engaged with the students during the module implementation in the class sessions. However, members of the research team were available inside the Minecraft server to assist in resolving potential student problems. The research of Manahan and Rodrigo (2022) provides a more thorough explanation of the preparation and support given to partner teachers and their classes in integrating and implementing WHIMC in their curriculum.

In this study, the partner teachers from a middle school in the Philippines developed two (2) learning modules for their Grade 8 science curriculum. Since Minecraft uses a biome system and adopts representation of real-world animals (Ekaputra et al., 2013), the partner teachers utilized WHIMC to teach topics on ecology. The partner teachers chose ecosystem as the topic for Module 1 and biodiversity and evolution for Module 2. The developed modules employed asynchronous and synchronous teaching modalities. The learning modules implemented a self-discovery teaching strategy where students are provided access to the WHIMC worlds before the 1 h synchronous sessions to give students ample time to explore, provide observations, and infer an understanding of the worlds. The Minecraft game-play was integrated into the modules as a pre-lecture and motivation activity. Wang et al. (2022) found that students at different educational levels respond differently to games. Primary school students are at a developmental stage where they are unable to master the rules of the games quickly and are therefore attracted by the freshness and novelty of games. However, secondary and higher education students master the game rules quickly, resulting in decreased interest. Thus, the Minecraft game-play integrated into the module has no specific time limit to allow students to explore the worlds at their own pace. However, each Minecraft session must be completed before the synchronous session. Students need to complete 2 Minecraft game-play sessions. The learning tasks integrated into the WHIMC-based modules were designed to apply a number of higher-order thinking skills represented in Bloom's taxonomy. The game attribute of the modules consists of the exploration of the simulated environment of the WHIMC worlds. Students underwent training and orientation in Module 1, wherein they explored the space station and experienced the hub that supports life. They explored the built-in ecosystem of the Lunar Base LeGuin to identify the biotic and abiotic components and observe the systemic relationships of the staff in the area. In Module 2, students explored the What-If worlds, wherein they experienced the life of an astronaut. They also experienced different What-If scenarios of the planet Earth (Tilted Earth, No Moon, Colder Sun) that showed them opportunities to observe the planet under altered conditions. The observation of the students must revolve around the environmental change of the different versions of Earth compared to normal Earth, the appearance of trees, plants, and topography, the existence and behavior of animals, and compare the pressure, temperature, oxygen, radiation, atmosphere, altitude, and wind for each world.

Each module began with an asynchronous session in which students explored the WHIMC worlds and recorded their observations as indicated in the module. After the asynchronous session, students turn in their answers for the formative assessments and activity worksheets. The 1 h synchronous session focused on the discussion of the lesson using simulations and inquiry-based learning to encourage student active participation, followed by a knowledge assessment related to the topic. See Figures 2, 3 for the excerpt of the developed WHIMC-based modules.

2.2. Participants

The entire Grade 8 school population consisting of 8 class sections were recruited for the study. However, out of the 212 prospective participants, 31 opt not to participate and 64 did not complete the survey questionnaires they were asked to answer. Thus, the total participants in this study were 117 middle school students (53 male and 64 female) aged 13–14 years old. The collection of data from the participants was approved by the University Research Ethics Office (UREO). The students submit the signed consent forms

Module 1 Topic: Ecosystem

Learning objectives:

- Identify the biotic and abiotic components of the environment
- Suggest ways and possible individual or collective solutions to minimize human impacts on the environment

Asynchronous Session:

- Exploration of Minecraft WHIMC World
    - Lunar Base LeGuin
- Learning Task
    - Observe the mini ecosystem in the Biodome and identify the biotic and abiotic components
    - Observe the systemic relationship of astronauts, scientists, engineers, and staff to keep the world working
    - Complete the activity worksheet by identifying the effects of each abiotic factor in the worksheet and cite at least one example

Synchronous Session:

- Students share their experience in exploring the Minecraft WHIMC world
- Introduction of the lesson or topic about the ecosystem and the organisms that lived on it
- Discussion of the lesson by showing a video of the simulated environment through the Minecraft WHIMC world and asking students targeted guide questions to facilitate inquiry-based learning
- Assessment
    - Answer the identification questions related to the topic
    - Differentiate biotic and abiotic factors and give an example of each king of factor
    - Answer the essay question: How can human activities affect the balance of ecosystems

FIGURE 2
Lesson excerpt of module 1.

indicating their participation in the experiment prior to data collection. The data used in the analysis come from the Stem Interest Questionnaire (SIQ) ratings, Game Experience Questionnaire (GEQ) ratings, and answers to the open-ended questions about the module attributes, alongside the performance ratings (high or low) of the participants.

2.3. Pre-test and post-test

Before using WHIMC, the students complete the pre-SIQ to determine their baseline interest in these domains. The students took knowledge assessments, the GEQ, and the post-SIQ as post-test after using WHIMC. The SIQ was given as a pre-test and post-test to determine whether using the WHIMC-based modules made an impact on the STEM interests of students.

2.4. Knowledge assessment

Students took knowledge assessments after the asynchronous and synchronous sessions of each module. The out-of-game assessments consisted of formative evaluations, asynchronous worksheets, and long tests. The observations made by the students while using WHIMC served as formative evaluations. After the asynchronous session, students must complete the asynchronous worksheets associated with each module topic. Further, long tests consisting of identification and essay questions related to the module topics were administered after the synchronous sessions. High-performers and low-performers were identified based on their out-of-game assessment scores. High-performers (HP) are those students with total assessment scores exceeding the mean score ($HP = s > \bar{x}$). Conversely, low-performers (LP) are those students with total assessment scores below or equal to the mean score ($LP = s \leq \bar{x}$).

Module 2 Topic: Biodiversity and Evolution: Adaptation

Learning objectives:

- Infer the possible adaptations that the organisms must undergo based on your observation in the What If world.
- Explain how this adaptation can lead to species diversity and survival.

Asynchronous Session:

- Exploration of Minecraft WHIMC World
  - Tilted Earth
  - No Moon
  - Colder Sun
- Learning Task
  - Observe the environmental change of the Tilted Earth, No Moon, and Colder Sun compared to normal Earth
  - Observe the appearance of tress, plants, and topography of the What If worlds
  - Observe the existence and behavior of animals in the What If worlds
  - For each world, check the following factors and take note for comparison: pressure, temperature, oxygen, radiation, atmosphere, altitude, wind
  - Complete the activity worksheet by creating a concept map using the terms provided in the worksheet. Start with the word Evolution.

Synchronous Session:

- Students share their experience in exploring the Minecraft WHIMC world
- Introduction of the lesson or topic about biodiversity and evolution
- Discussion of the lesson by showing pictures of extinct animals, showing videos of the simulated environment of Tilted Earth, No Moon, and Colder Sun worlds and asking students targeted guide questions to facilitate inquiry-based learning
- Assessment
  - Answer the following essay questions
  - As people moved from place to place, they have often brought plants and animals with them. How might the introduction of a new species of plant and animal in an area have disastrous effects on organisms already in that area?
  - How is biodiversity important in an organism's survival?

FIGURE 3
Lesson excerpt of module 2.

2.5. Stem interest questionnaire

The pre-SIQ determined their interests prior to using WHIMC. After answering the SIQ, students were given access to the WHIMC worlds and instructed to follow the guidelines described in the teacher-created learning modules. Students then answered the post-SIQ and the Game Experience Questionnaire (GEQ) after using WHIMC. The out-of-game assessment questions that are part of the teacher-created learning modules were then given to the students to complete the data collection process.

The SIQ used in this study is an abridged version of an original Student Interest and Choice in STEM (SIC-STEM) questionnaire developed by Roller et al. (2018), which was based on the Social Cognitive Career Theory (SCCT) questionnaire of Lent and Brown (2008). This instrument is employed to characterize and assess the propensity of students to pursue STEM careers. In this framework, five dimensions (SCCT constructs) are identified to describe STEM interests: *Self-efficacy*: the judgment of one's perceived ability; *Outcome Expectations*: the perceived consequences of one's decisions and; *Interests*: the affinities of a person; *Choice Goals*: the perception that the choice to acquire STEM-related knowledge is important in the future; and *Choice Actions*: the perception that STEM-related actions today will provide support in a future career.

The SIQ used in this study consisted of 10 items from the SIC-STEM questionnaire based on their relevance to WHIMC and the teacher-created learning modules. The respondents rate their level of agreement using a 5-point Likert scale format (1 – *strongly disagree*, 2 – *disagree*, 3 – *neutral*, 4 – *agree*, 5 – *strongly agree*). Table 1 presents the mapping of the SIQ items to the SIC-STEM constructs.

2.6. Game experience questionnaire

The GEQ used in this study is also an abridged version of the instrument developed by IJsselsteijn et al. (2013) to measure the factors in a game that contribute to an engaging *gameful* experience described across seven (7) dimensions of the player experience namely, *Immersion*: how strongly the players felt connected to the game; *Flow*: how much the player lost track of their own effort or time while playing the game; *Competence*: the player's judgment of their own performance against the game's goals; *Positive Affect* and *Negative Affect*: reports of positive and negative emotional experiences while playing the game; *Tension*: reports relating to frustration and annoyance; and *Challenge*: an indication of how difficult the players found the game to be. Johnson et al. (2018) validated the GEQ used in this study and the findings suggest a revised structure that reduces the seven dimensions to five factors. *Flow, immersion, competence*, and *positive affect* dimensions have some empirical support. However, it was noted that items in the *negative affect, tension*, and *challenge* dimensions overlap and should not be evaluated independently. It would be more acceptable to see these aspects as being merged into a single negative factor. Since we wanted a fine-grained analysis of the negative gaming experience of the students while using WHIMC, we treated the *negative affect, tension*, and *challenge* dimensions separately.

The questionnaire used in this study was adopted from Casano and Rodrigo (2022b). The instrument only included 23 items that seemed relevant to the context of WHIMC and the teacher-created learning modules out of the 33 core module items of the original GEQ. The respondents rate their level of agreement with the items using a 5-point Likert scale format (*not at all - 1, slightly - 2, moderately - 3, fairly - 4, extremely - 5*). Table 2 presents a mapping of the GEQ items to the player experience components.

TABLE 1 Mapping of SIQ items to the SCCT constructs.

SIC-STEM constructs	Items
(**SE**) Self-Efficacy	1 I know I can do well in science.
	4 I think Science is challenging to learn.
(**OE**) Outcome Expectations	9 After I finish high school, I will use Science often.
	10 I believe that I can use Math and Science to solve problems in the future.
(**I**) Interests	2 I enjoy Science activities.
	3 I enjoy solving Science and Math problems.
(**CG**) Choice Goals	5 Leaning Science will help me get a good job.
	6 Knowing how to use Math and Science together will help me to invent useful things.
	7 Understanding engineering is not important for my career.
(**CA**) Choice Actions	8 I try to get a good grade in science because I have an interest in science jobs.

Four (4) open-ended questions were appended to the GEQ. These questions were: *What was your favorite part of the module and why?; What was your least favorite part of the module and why?*; *What about WHIMC made the topic fun, interesting, or easy to learn?;* and *What about WHIMC made the topic boring and/or difficult to learn?.*

2.7. Data analysis

To answer the research questions of this study, we conducted statistical analyses of the pre-SIQ and post-SIQ, GEQ, and answers to the open-ended questions on the module attributes. Paired samples *t*-test was used to analyze the pre-SIQ and post-SIQ ratings of the students to determine the effect of using WHIMC on the STEM interests of students. Independent samples *t*-test was used to compare the game experience between the high-performers and low-performers using their GEQ ratings. A point-biserial correlation was used to determine the strength and direction of association of each favorite module attribute between the high-performers and low-performers.

For the qualitative analysis, the text data (responses to the favorite and least favorite open-ended questions on module attributes) were analyzed using thematic analysis. The text data were assessed and tagged by coders as being related to the learning topic, learning task, or game attribute of the teacher-created learning module.

The resulting dataset was then subjected to the *bag-of-words* approach for text analytics. In particular, pre-processing was conducted to transform the text data into a quantifiable form. The text data was converted into lowercase form, removal of punctuations, special symbols, numbers, and extra whitespaces, *stopwords* (pronouns and other common yet irrelevant words), stemming (transformation to base form), and stem completion (transformation to sensible form). Finally, the text data were tokenized and transformed into a document-term matrix.

The transformed text data was then merged with the performance and thematic tagging data, and were then subjected to statistical treatments. Descriptive visualizations were employed to characterize the responses of the students. Word clouds were used to show the relative frequencies of dominant words for each module and each type of response (favorite or least favorite attribute).

3. Results

3.1. Analysis of SIQ ratings

The students answered the SIQ twice: before and after playing WHIMC. A paired-samples *t*-test was conducted to compare the SIQ ratings of the students before and after using WHIMC as a learning tool. The analysis of the SIQ ratings revealed that there was no significant difference in the overall pre-SIQ ratings ($M = 3.60$, $SD = 0.27$) and post-SIQ ratings ($M = 3.65$, $SD = 0.29$) using WHIMC; $t(116) = -1.78$, $p = 0.077$. There is only a slight increase in the overall SIQ ratings after using WHIMC. This result suggests that using WHIMC as a learning tool only has a minimal effect on the STEM interests of the students.

To conduct further analysis on the SIQ ratings, paired samples t-tests were conducted to compare the SIQ ratings of the students

TABLE 2 Mapping of GEQ items to the player experience components.

GEQ component	Items	GEQ component	Items
(**I**) Immersion	2 I was interested in the game's story	(**P**) Positive Affect	1 I felt content.
	9 It was esthetically pleasing.		3 I thought it was fun.
	14 I felt imaginative.		5 I felt happy.
	15 I felt that I could explore things.		10 It felt good.
	19 I found it impressive.		
	22 It felt like a rich experience.		
(**F**) Flow	4 I was fully occupied with the game.	(**N**) Negative Affect	6 It gave me a bad mood.
	20 I was deeply concentrated on the game.		7 I found it tiresome.
			12 I felt bored.
(**C**) Competence	8 I felt competent.	(**T**) ension	17 I felt annoyed
	11 I was good at it.		21 I felt frustrated
	13 I felt successful.		
	16 I was fast at reaching the game's targets.		
		(**CH**) Challenge	18 I felt challenged
			23 I felt time pressured

before and after using WHIMC on the different SIC-STEM constructs. The result of the statistical analysis revealed that only the Choice Actions construct of the 5 SIC-STEM constructs showed a statistically significant difference. The pre-SIQ rating of the Choice Actions construct ($M = 3.34$, SD = 1.13) significantly increased after using WHIMC ($M = 3.50$, SD = 1.00); $t(116) = -2.263$, $p = 0.025$. This result indicates that the students understood the importance of studying hard and earning high marks in class if they are interested in STEM-related careers. Figure 4 presents the bar chart showing the aggregated pre-SIQ and post-SIQ ratings on each SIC-STEM construct.

Figure 5A shows the bar charts of the pre-SIQ and post-SIQ ratings on each SIC-STEM construct of the low-performers. Paired-samples t-tests were conducted on each construct and results show that the pre-SIQ rating for the Self-efficacy construct ($M = 3.49$, SD = 0.59) significantly increased after using WHIMC ($M = 3.63$, SD = 0.67); $t(53) = -2.127$, $p = 0.038$. This finding might indicate that the low-performers gained some confidence in their ability to understand science concepts.

Figure 5B shows the bar charts of the pre-SIQ and post-SIQ ratings on each SIC-STEM construct of the high-performers. Paired-samples t-tests were conducted on each construct and results revealed that the pre-SIQ rating for the Interest construct ($M = 3.60$, SD = 0.77) significantly increased after using WHIMC ($M = 3.74$, SD = 0.80); $t(62) = -2.092$, $p = 0.041$. High-performers' increased level of agreement in the Interests construct may be related to how much they enjoyed and persisted in completing the assigned tasks from the WHIMC-based modules.

The observations on the analysis of each SIC-STEM construct provided some evidence that the teacher-created learning modules using WHIMC increased some aspects of STEM interest among students.

3.2. Analysis of the GEQ answers

The GEQ was administered to measure the factors in a game that contribute to an *engaging gameful experience* described across 7 dimensions of the player experience: Positive Affect (PA), Negative Affect (NA), Immersion (I), Flow (F), Competence (C), Challenge (Ch), and Tension (T). Independent samples t-test was used to determine if there is a significant difference in the overall GEQ ratings between the high- and low-performers. The statistical test result revealed no statistically significant difference in the overall GEQ ratings between the high-performers ($M = 2.48$, SD = 0.13) and low-performers ($M = 2.38$, SD = 0.19); $t(103) = -1.311$, $p = 0.193$. This result revealed that both groups had the same level of engagement in using WHIMC as a learning tool. Independent samples *t*-tests were used on each dimension to check for differences between high- and low-performers. The tests revealed that only the Immersion dimension had a significant difference between the groups. High-performers have significantly higher GEQ ratings ($M = 3.34$, SD = 0.56) compared to the low-performers ($M = 3.04$, SD = 0.68) after using WHIMC; $t(106) = -2.584$, $p = 0.011$. This finding suggested that high-performers connected more deeply with the game and may therefore have had a more engaging learning experience than low-performers. Figure 6 shows the GEQ ratings of the high- and low-performers on each GEQ dimension.

3.3. Analysis of the open-ended answers

Insights drawn from analyzing the answers to the open-ended questions about the module attributes might complement the observations from the analysis of the SIQ and GEQ ratings discussed in the previous sections. We conducted qualitative analysis of the responses to the open-ended questions to determine the relationship between the module attributes and student performance.

The individual answers of the students about their favorite and least favorite attributes of the module were assessed and tagged as feedback about the learning topic, learning task, or game module attribute. Three coders categorized 468 rows of open-ended answers using the criteria described in Table 3. The coders coded independently using a spreadsheet containing the class numbers with the

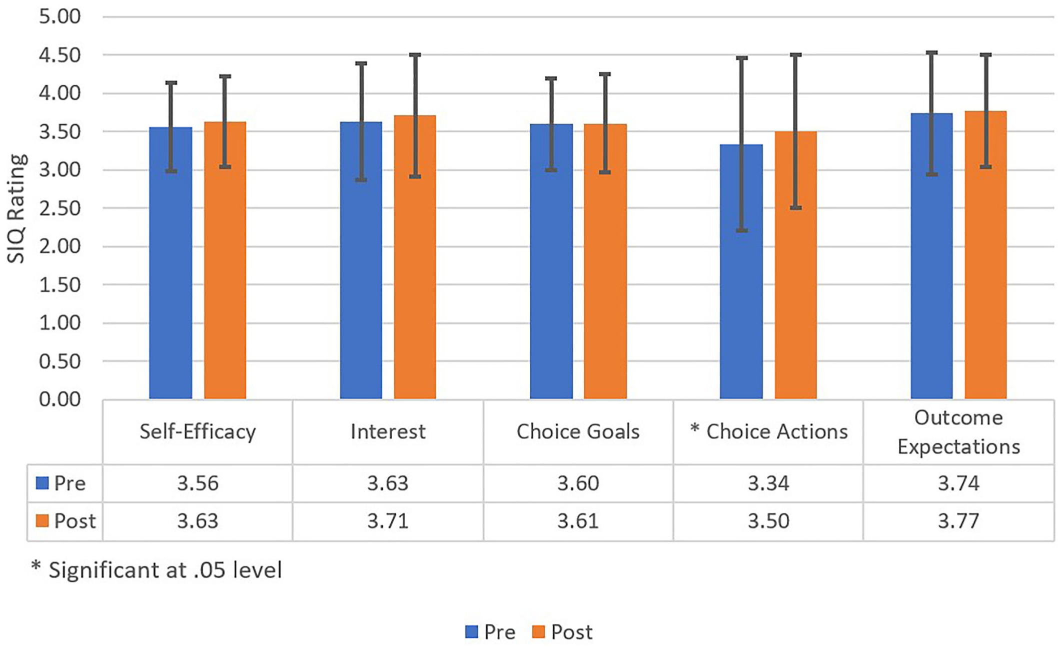

FIGURE 4
Aggregated pre- and post-SIQ ratings on each SIC-STEM construct.

corresponding open-ended answers and three (3) columns with headings indicating the three module attributes. Each coder tagged the open-ended answer by filling in the columns with either 1 or 0 indicating the presence or absence of the module attribute in the feedback. The coders unanimously coded 995 (70.87%) module attributes, two (2) coders were in agreement for the 383 (27.28%) module attributes, and 26 (1.85%) module attributes were coded differently by each coder. The coders then convened to reach a consensus on the differences in the coding.

3.3.1. Analysis of the answers to the favorite part of the module

The 234 rows of labeled data containing the values of favorite module attributes were analyzed using frequency count to determine the favorite module attributes and the number of favorite attributes. A point-biserial correlation was also performed to determine the strength and direction of association of each favorite module attribute between the high-performers and low-performers. This statistical analysis was utilized since the nature of the data is dichotomous.

The *bag-of-words* text analytics approach was then applied to the text data. The transformed text data was then merged with the performance for quantitative text analytics. This analysis was performed to characterize the text data and identify the underlying themes.

Figure 7A shows that the favorite module attribute of both groups is the learning topic of the modules. This result implies that high-performers and low-performers enjoyed the lessons integrated into the WHIMC-based learning modules. High-performers liked all the module attributes except the learning task attribute of Module 2. On the other hand, low-performers prefer the learning topic module attribute over the learning task and game module attributes.

The percentage of respondents on the number of favorite attributes (Figure 7B) revealed that most of the low performers mentioned 2 module attributes whereas high performers mentioned 3 module attributes in their responses about their favorite attributes in Module 1. However, for Module 2, both groups identified only one (1) module attribute as their favorite. Based on the data presented in Figure 7A, low-performers chose the learning topic and tasks as their favorite module attributes of Module 1. Further, both groups liked the learning topic more than the learning task and game module attributes of Module 2.

Table 4 presents the point-biserial correlation result of the favorite module attributes. The table shows a significant positive correlation between the **game** module attribute and performance ($rpb = 0.203$, $n = 117$, $p = 0.029$). This implied that students who liked the game attribute of Module 1 performed better in the out-of-game assessments. For Module 2, the performance has significant positive correlation with the **learning task** ($rpb = 0.270$, $n = 117$, $p = 0.003$) and **game** ($rpb = 0.307$, $n = 117$, $p = 0.001$) module attributes while a significant negative correlation was observed for the **learning topic** ($rpb = -0.237$, $n = 117$, $p = 0.010$). This finding could mean that students who chose the learning topic module attribute as their favorite did not perform well in the assessment. In contrast, students who performed better in the assessment chose the game or learning task module attribute as their favorite part of the module. We also found a significant positive correlation between the number of favorite attributes of Module 1 ($rpb = 0.208$, $n = 117$, $p = 0.024$) and Module 2 ($rpb = 0.212$, $n = 117$, $p = 0.022$) with the performance.

These findings corroborate the result of the analysis of the GEQ ratings that high performers had a better quality of game experience compared to low performers. Students who liked the game and learning task module attributes are likely to perform better in the out-of-game assessments. We note that 2 out of the 3 out-of-game assessments are conducted after exploring the WHIMC worlds assigned in the modules. Thus, students must be engaged in the game and learning tasks to have better assessment scores.

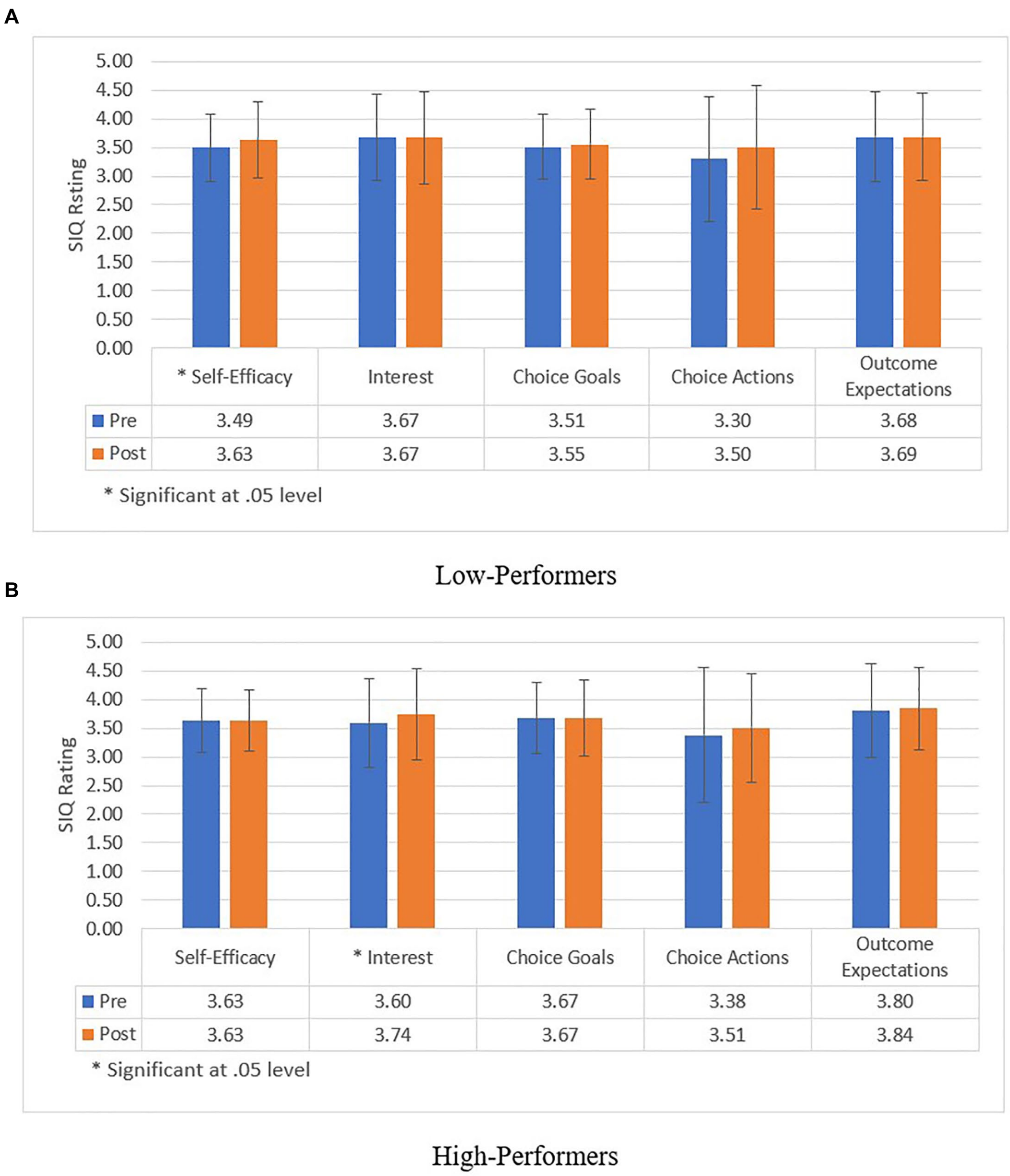

FIGURE 5
Pre- and post-SIQ ratings on each SIC-STEM construct, **(A)** Low-performers. **(B)** High-performers.

To characterize the responses of the high- and low-performers to the open-ended questions, word clouds were generated. As can be seen in Figure 8A, the most dominant word about the favorite attribute of Module 1 is *learn*. This finding suggests that both high performers and low performers mentioned learning in their responses. The other dominant words such as *Minecraft* and *fun* refer to the simulated environment using WHIMC, which is related to the game attribute of the module. The words *ecosystem*, *biotic*, and *abiotic* are related to the topic or lessons in Module 1. The word *explore* might be related to the learning task module attribute since students were asked to explore the WHIMC world Lunar Base LeGuin to identify the biotic and abiotic components and make observations about the systemic relationships of the people. This finding is aligned with the results of the quantitative analysis of the tagged text data since the dominant words relate to all the module attributes.

Similar to the findings in the responses about the favorite attributes of Module 1, *learn* is also the top word in the responses about the favorite attribute of Module 2 (Figure 8B). The words *different*, *worlds*, *explore*, and *fun* might refer to the ability of the students to explore the different worlds and the fun experience they had using WHIMC. These words are related to the game attribute of the module. The words that relate to the learning topic attribute are *animals*, *things*, *interesting*, and *adapt*. Students did not mention much in their responses about *quests* and *observations*, which are words related to the learning task attribute. This result indicates that while the students enjoyed the learning topic and game component of Module 2, they were less enthusiastic about the learning tasks.

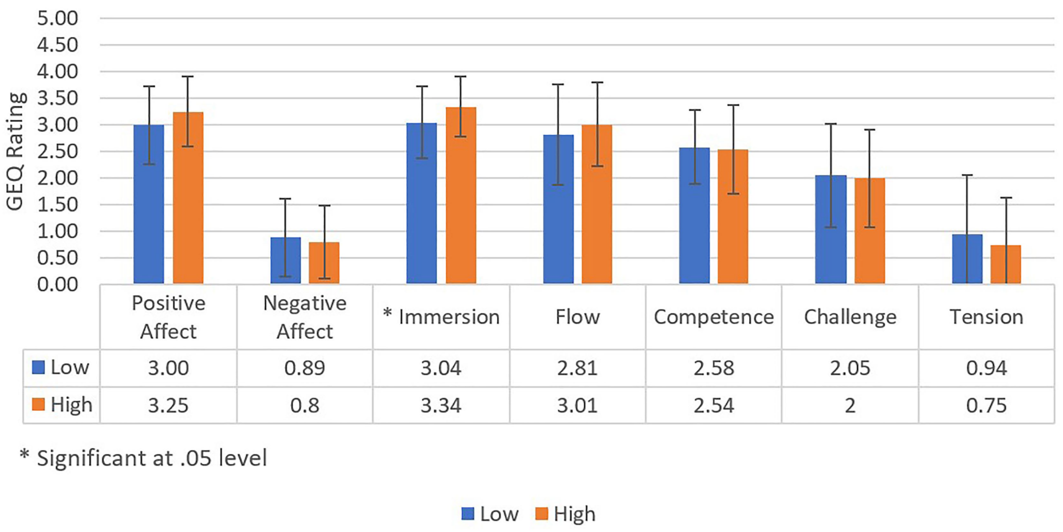

FIGURE 6
Game experience dimensions between the high- and low-performers.

TABLE 3 Attributes of the teacher-created learning modules.

Module attribute	Criteria
Game	If the answer mentions elements of the WHIMC map or interactions within the game world including references to in-game mechanics, the answer is categorized as *Game*.
Learning Topic	If the answer mentions being able to acquire information in some way, or learning facts while interacting with the WHIMC worlds, the answer is categorized as *Learning Topic*.
Learning Task	If the answer makes a reference to the tasks or mentions an in-game behavior as indicated in the teacher-created learning module, tag the answer with *Learning Task*.

3.3.2. Analysis of the answers to the least favorite part of the module

The same analysis discussed in the analysis of the answers to the favorite part of the module was also utilized to draw insights about the least favorite part of the module.

Based on Figure 9A, the game and learning task attributes of Module 1 are the least favorite. This result might be because students encountered technical difficulties while playing and experienced a hard time completing the quests or tasks assigned in the module. For Module 2, most of the comments come from the high-performers and they identified the learning task module attribute as their least favorite. This might be because of the many tasks assigned in this module and the need to go through 3 What-If worlds, which require more time to complete and more observations to be recorded while playing the game.

Figure 9B presents the number of least favorite attributes of the high- and low-performers. We can observe that at least 1 module attribute has been mentioned by both groups. The game attribute of Module 1 as shown in Figure 9A was identified to be the least favorite of both groups. However, for Module 2, most of the low-performers did not have a least favorite whereas high-performers mentioned at least one least favorite module attribute. The high-performers are less enthusiastic about the learning task module attribute.

The result of the point-biserial correlation shows that the attributes of Module 1 and the number of least favorite attributes have no significant correlation with student performance as shown in Table 5. This result could mean that although students mentioned attributes of the module that they do not like, it does not influence their performance. In terms of Module 2, the Task module attribute has a significant positive correlation with student performance ($rpb = 0.327$, $n = 117$, $p = <0.001$) and the number of favorite attributes ($rpb = 0.202$, $n = 117$, $p = 0.029$). The result implies that students who mentioned the Task module attribute as their least favorite perform better than those who did not. When high-performers comment about the learning task module attribute, this might be because they experienced a hard time doing the assigned tasks but are still motivated to complete them.

To characterize the responses of the high- and low-performers to the open-ended questions on the least favorite module attributes, word clouds were generated. The top five dominant words for the responses on the least favorite attributes of Module 1 (Figure 10A) are *time*, *Minecraft*, *hard*, *going*, and *confusing*. These words describe the experience that the students had while playing WHIMC. Students mentioned in their comments that they had a hard time connecting to Minecraft, going to different worlds or portals, and sometimes being confused about what to do next. This finding implies that most of the comments are related to the game attribute of the module.

The top five dominant words for the responses on the least favorite attributes of Module 2 (Figure 10B) are *time*, *quests*, *Minecraft*, *confused*, and *find*. These words relate to the experience that the students had while doing the tasks integrated into the module using WHIMC. Students commented about experiencing a hard time

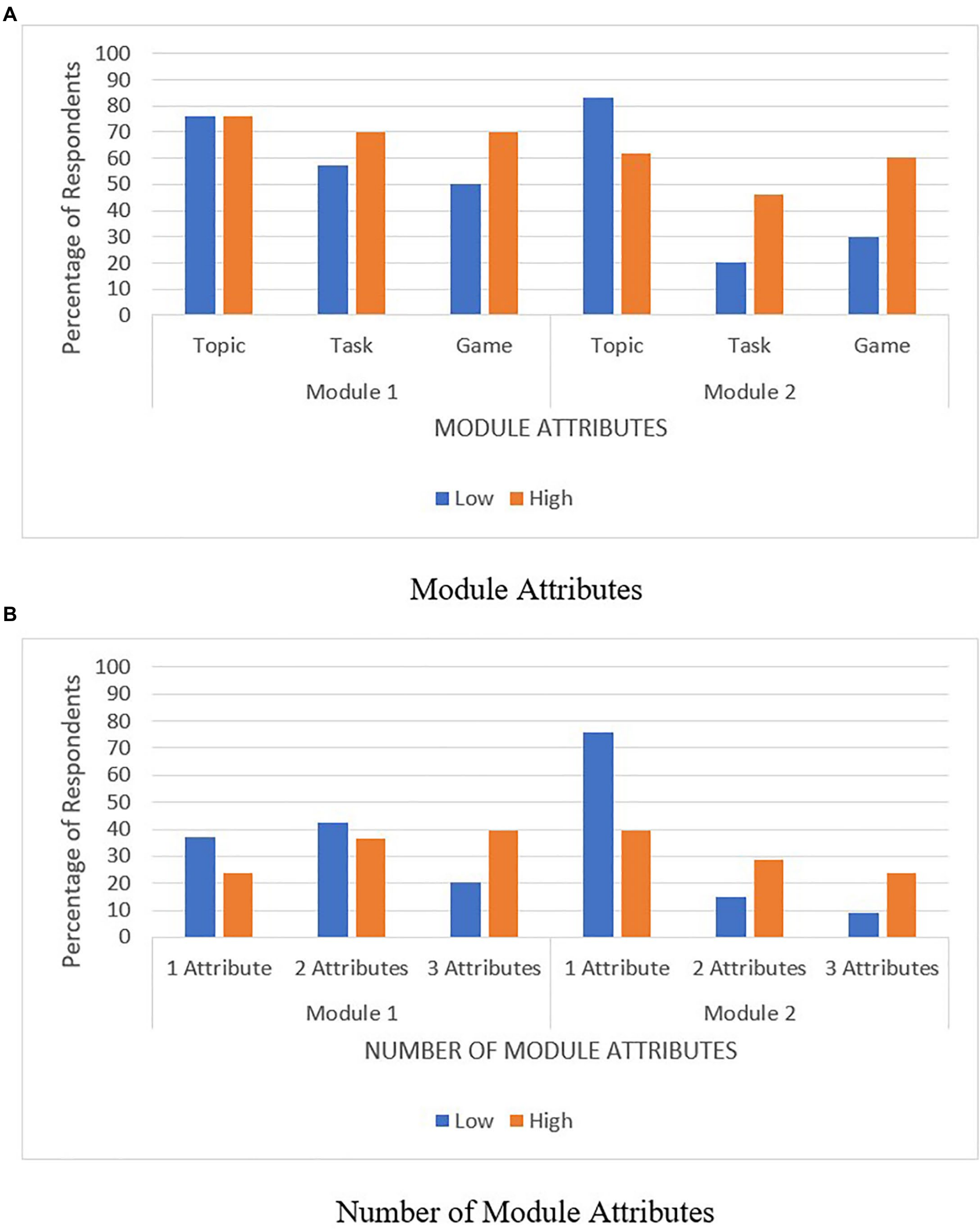

FIGURE 7
Responses on favorite module attributes, **(A)** Module attributes, **(B)** Number of module attributes.

TABLE 4 Point-biserial correlation result of the favorite module attributes.

Variables	Statistics	Topic	Task	Game	No. of favorite attributes	Topic	Task	Game	No. of favorite attributes
		Module 1				Module 2			
Performance	Point Biserial	0.003	0.129	0.203*	0.208*	−0.237**	0.270**	0.307**	0.212*
	Sig. (2-tailed)	0.974	0.165	0.029	0.024	0.010	0.003	0.001	0.022

completing the quests, finding the NPCs, and being confused about where to go next to complete the quests. These comments relate to the task and game attributes of the module.

Why did student preferences differ from Module 1 to Module 2? We offer some speculation: The learning objectives of Module 1 were simple (see Figure 2), and students only had to explore the biodome

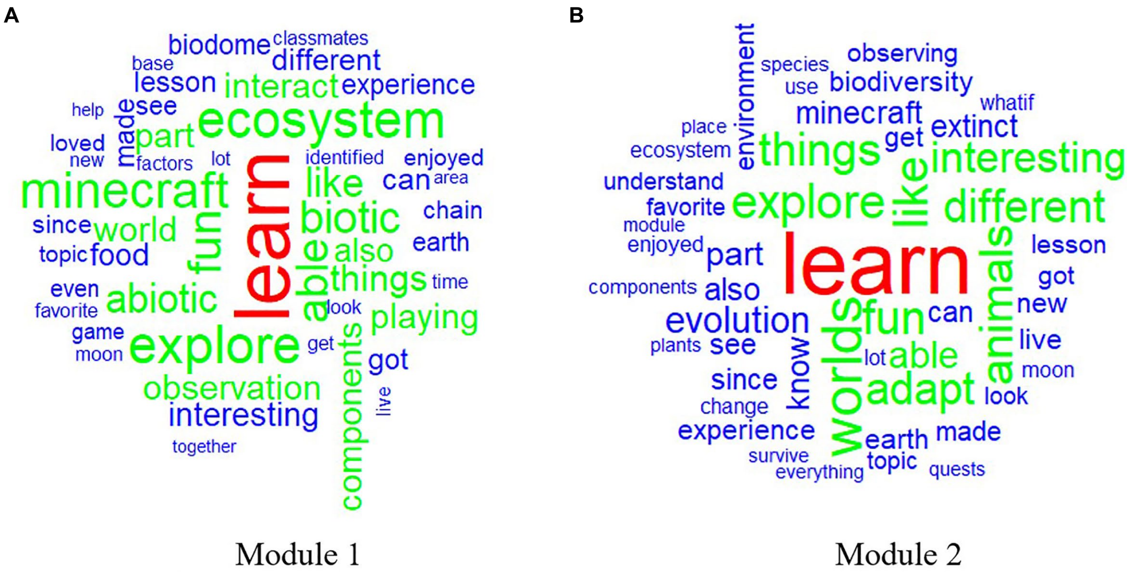

FIGURE 8
Frequencies of dominant words on the favorite attributes. **(A)** Module 1, **(B)** Module 2.

to perform the learning tasks and get the answers to the out-of-game assessments. This means that Module 1 tended to be easy, which may account for why high-performers liked all the module attributes and low-performers liked the topic and task attributes. Low-performers did not express liking the game attribute, a sentiment echoed by their GEQ responses, in which they had slightly higher ratings for Negative Affect and Tension dimensions compared to high-performers. Low-performers might have found the open-ended learning environment confusing. They might not have had a high-level understanding of their location, leading them to wander without purpose (Esclamado and Rodrigo, 2022a).

For Module 2, students had to explore three WHIMC worlds (Tilted Earth, No Moon, Colder Sun). They had to make observations to infer the possible adaptations of organisms and explain how these adaptations could lead to species diversity and survival. Module 2 was harder and more open-ended than Module 1. This might explain why many high performers expressed not liking the task module attribute.

4. Discussion

Learners often find STEM difficult because it requires complex thinking, repeated practice, and self-discipline. Hence, educators are thinking of innovative ways to provide an engaging learning environment that keeps students interested and enthusiastic about STEM subjects. Minecraft is one of the innovative approaches that has been adopted in science education (Pusey and Pusey, 2015; Nkadimeng and Ankiewicz, 2022). Thus, the purpose of this study is to continue to promote the use of WHIMC-based modules as a learning tool to cultivate the STEM interests of Filipino middle school students.

Our first research question is to determine the effect of using WHIMC on the STEM interests of students. The result of the analysis of the aggregated SIQ ratings before and after using the WHIMC-based modules revealed only a minimal effect on the STEM interests of the students. This implies that the implementation of the WHIMC-based modules in a Philippine middle school did not reveal a significant impact on the students' STEM interests based on their SIQ ratings. This finding supports the result of the analysis of the STEM interest of primary school learners in the Philippines (Casano and Rodrigo, 2022b). But the result of this study is promising since there is still an increase in the SIQ ratings of students after learning two ecology topics with WHIMC. Further, there is a significant increase in the Choice Actions construct, which suggests that the students appreciate the importance of motivation to study hard and get good grades if they want to pursue STEM-related careers. Moreover, the significant increase in the Self-efficacy ratings of low-performers might suggest that they gained some confidence in their ability to understand science concepts after using WHIMC. This result is aligned with Nkadimeng and Ankiewicz (2022) that using MinecraftEdu helped students gain a more concrete understanding of abstract topics. High-performers' increased level of agreement in the Interests construct may be related to how much they enjoyed and persisted in completing the assigned tasks from the WHIMC-based modules. This result corroborates the findings that using a digital game in teaching may be successful in fostering STEM interest (Bonner and Dorneich, 2016; Saricam and Yildirim, 2021; Ishak et al., 2022). Development of additional WHIMC-based modules focused on ecology topics might be needed to conduct a further evaluation to confirm or contrast the result of this study. This endeavor will be challenging since successful module design and implementation is time-consuming and requires technical, pedagogical, and content knowledge.

We also wanted to find out if there is a difference in the game experience between the high-performers and low-performers. The analysis of the overall GEQ ratings revealed no statistically significant difference between the game experience of high- and low-performers. We can infer that the overall game experience with WHIMC was the same for high and low performers, which confirms the finding of Casano and Rodrigo (2022b). Statistical tests were also conducted for

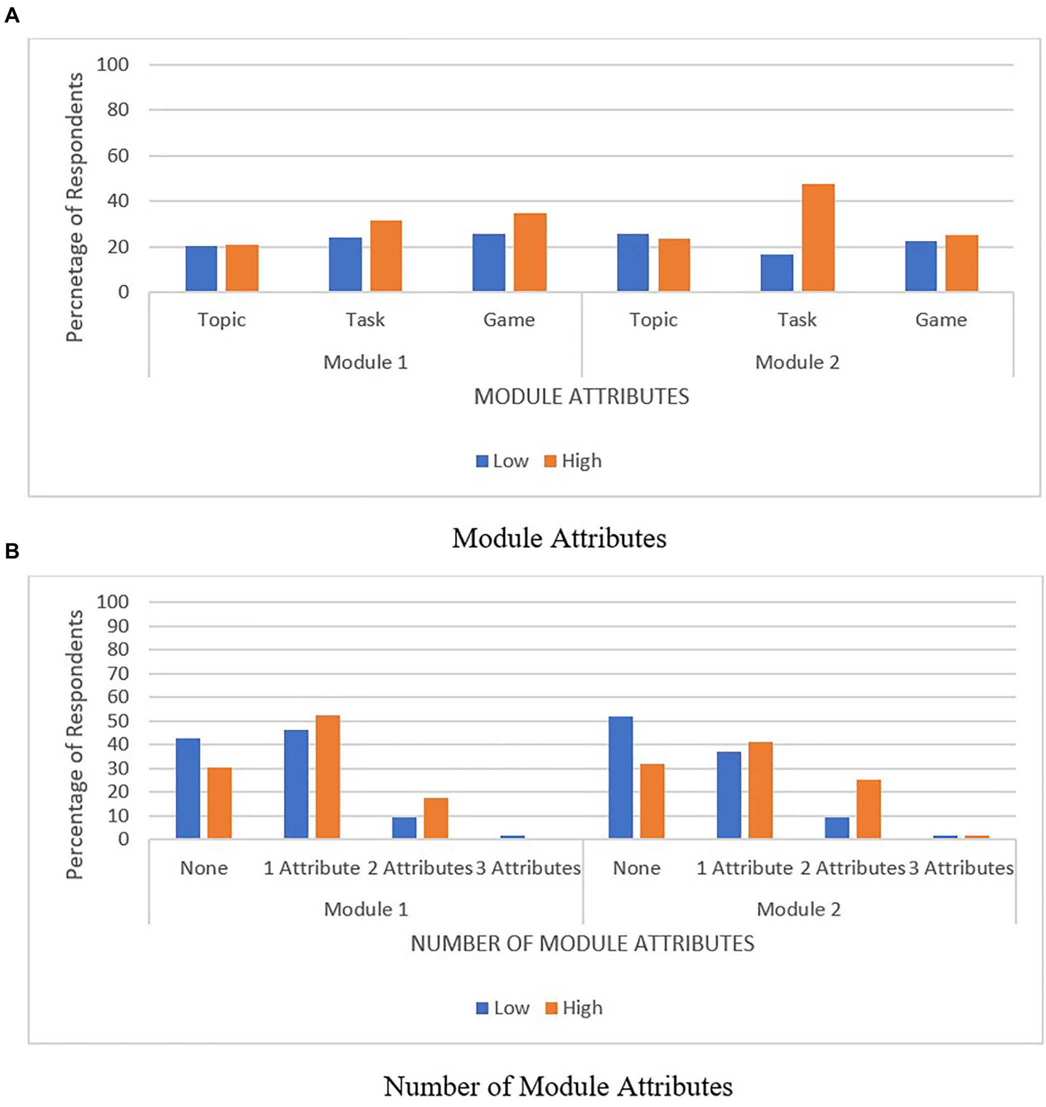

FIGURE 9
Responses on Least Favorite Module Attributes. **(A)** Module Attributes, **(B)** Number of Module Attributes.

TABLE 5 Point biserial correlation result of the least favorite module attributes.

Variables	Statistics	Topic	Task	Game	No. of favorite attributes	Topic	Task	Game	No. of favorite attributes
		Module 1				Module 2			
Performance	Point Biserial	0.003	0.085	0.097	0.0.121	−0.0.024	0.327**	0.037	0.202*
	Sig. (2-tailed)	0.972	0.362	0.297	0.194	0.794	0.000	0.691	0.029

each GEQ dimension to check if there were dimensions that would reveal statistical significance between the high- and low-performers. Among the 7 GEQ dimensions, only the Immersion dimension showed a statistically significant difference between the groups. Although they have the same level of agreement for Negative and Positive Affect, Challenge, Competence, Flow, and Tension, high-performers have significantly higher GEQ ratings on the Immersion dimension. With this finding, we can infer that high-performers had a more positive, engaging, and enjoyable learning experience with WHIMC than the low-performers. These results support the findings of other studies that game-based learning could increase learning achievement (Hwang et al., 2012; Chu and Chang, 2014), engagement (Lester et al., 2014; Alonso-Fernandez et al., 2019; Gadbury and Lane, 2022), desire to learn (Gadbury & Lane), and enjoyment (Alawajee, 2021).

Lastly, we wanted to determine the relationship between the module attributes and student performance. The results of the thematic analysis of the open-ended questions revealed that the WHIMC-based module attributes could affect the student performance and interests of students in learning science concepts. The findings on the favorite

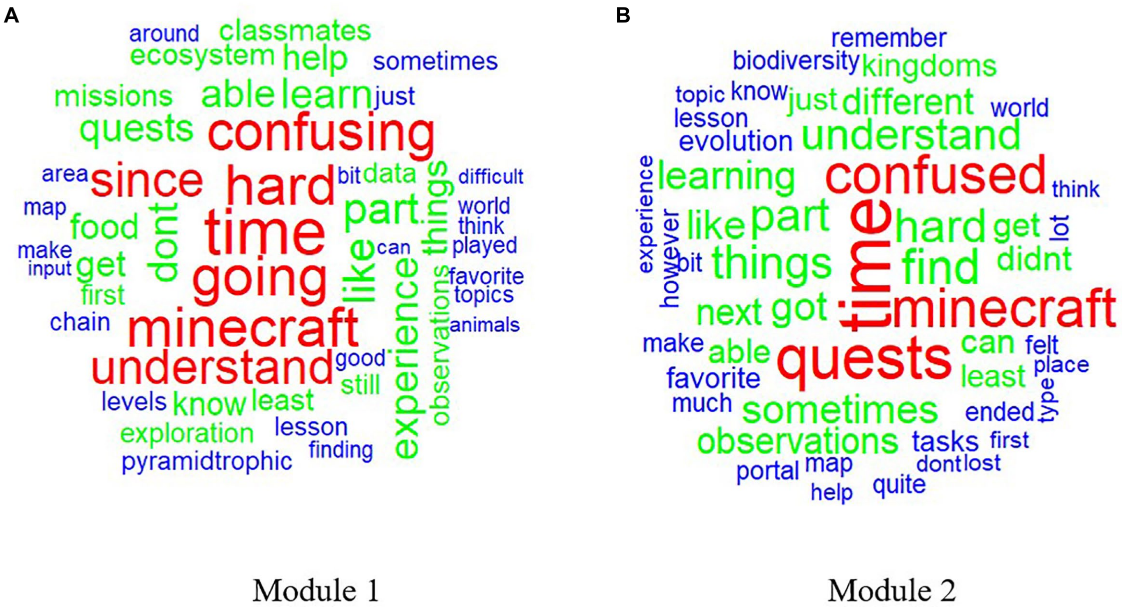

FIGURE 10
Frequencies of Dominant Words on the Least Favorite Attributes. **(A)** Module 1, **(B)** Module 2.

module attributes suggest that students perform better in the out-of-game assessments when they like all the module attributes. This implies that students must be engaged in the game and learning task aside from being interested in the learning topic to have better assessment scores. This finding corroborates the result of the analysis of GEQ ratings, where high-performers have higher ratings for immersion and flow dimensions after using WHIMC. The dominant words and themes of responses relate to the integration of WHIMC into the modules that allow students to learn and have a fun and enjoyable learning experience. The comments about the students' ability to understand the topics and the fun experience they had with the WHIMC-based modules could inform us about the suitability of using WHIMC as a learning tool in science education.

The findings on the thematic analysis of the least favorite module attribute revealed that the game and learning task attributes are the least favorite for Module 1. This result might be because students encountered technical difficulties while playing and experienced a hard time completing the quests or tasks assigned in the module. This result is aligned with the findings of Casano and Rodrigo (2022b) that low performers experienced difficulty in learning because of technical bugs and the learning tasks made it difficult for high performers to learn. For Module 2, most comments come from the high-performers who identified the learning task module attribute as their least favorite. This finding might be because of the many tasks assigned in this module, which require more time to complete and more observations to be recorded while playing the game. High-performers acknowledged the difficulty of the learning task but were still motivated to complete them. Students who did not cite any least favorite module attribute emphasized how fun learning was and how well they understood the lessons. The negative comments about the game and task attributes should be addressed in the future development of WHIMC-based modules to enhance the student learning experience and interests in STEM. Future module developments should consider the appropriate task completion duration since students can complete the tasks at different times. To alleviate the technical difficulties encountered while using WHIMC, partner teachers should organize more time for students to develop familiarity with the software so that they will be able to use the game's function effectively and efficiently.

The results of the thematic analyses on the favorite and least favorite module attribute are consistent with the findings about game-based learning. Researchers found that it improves student motivation (Hwang et al., 2012; Chu and Chang, 2014; Ennis, 2018; Leong et al., 2018; Hussein et al., 2019; Shang et al., 2019), encourages the player to learn (Iliya and Jabbar, 2015; Gadbury and Lane, 2022), helps in easy understanding of topics (Nkadimeng and Ankiewicz, 2022), and provides enjoyable coursework (Pusey and Pusey, 2015; Alawajee, 2021). With these findings, this research could contribute to the evidences of the impact of using game-based learning in teaching science concepts.

This research contributes to the literature in a number of ways. It suggests that an open-ended environment can be used to foster STEM interest, which corroborates previous findings on the use of Minecraft during summer camps (Yi et al., 2020, 2021; Lane et al., 2022). It collects and analyzes game-based data from the Philippines, a population that is underrepresented in the literature. It also contributes to the conversation about how and when games should be used with instruction. The study shows that Minecraft can be fun and engaging but just because it is fun and engaging does not guarantee that it will lead to increased interest in larger domains such as STEM. The study also shows that open-ended learning environments coupled with tasks that demand exploration, observation, and higher-ordered thinking are demanding even on high-performers. Low-performing students may require more scaffolding and guidance. Finally, the integration of educational games like Minecraft in classes requires lengthy lesson planning and technical preparation. Educators therefore have to curate the games well and monitor their outcomes in order to ascertain whether their use is truly worth the investment.

5. Limitations to the study

The work presented in this paper has some limitations. First, the analysis is only limited to the 2 WHIMC-based modules developed by partner teachers in a Philippine middle school. Thus, the findings from this initial study cannot be generalized because of the small number of topics used to determine the effect of using the modules on the STEM interests of students and game experience. We plan to have more partner teachers that will develop additional WHIMC-based modules and deploy these to other middle schools in the Philippines to see whether we can replicate the findings of this initial study.

During the module implementation, in-game data were also collected along with the SIQ, GEQ, and open-ended questions. So far, we have not yet analyzed the in-game data consisting of students' observations, use of science tools, and map explorations. In future work, we plan to analyze these in-game data to understand the in-game behaviors of students while interacting with the WHIMC worlds and their relationship to student performance and STEM interests.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by University Research Ethics Committee of the Ateneo de Manila University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

CT: writing, data analysis, coding, statistical analysis, and editing. MR: writing, editing, reviewing, and supervision. JC: writing and review. All authors contributed to the article and approved the submitted version.

Funding

Ateneo de Manila University and the Department of Science and Technology for the grant entitled, "Nurturing Interest in STEM among Filipino learners using Minecraft."

Acknowledgments

The authors thank H Chad Lane and Jeff Ginger for their enthusiastic collaboration, Dominique Maire Antoinette Manahan, Maricel Esclamado, Mikael William Fuentes, and Ma. Rosario Madjos for their support, our partner teachers and all participating students, Andhee Jacobe and Bobby Roaring for coding the module attributes, the Ateneo Laboratory for the Learning Sciences, and the Ateneo de Manila University. The authors thank our funding agency, the Department of Science and Technology for the grant entitled, "Nurturing Interest in STEM among Filipino learners using Minecraft," and our monitoring agency, the Philippine Council for Industry, Energy, and Emerging Technology Research and Development.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Alawajee, O. (2021). Minecraft in education benefits learning and social engagement. *Int. J. Game-Based Learn.* 11, 19–56. doi: 10.4018/IJGBL.2021100102

Alonso-Fernandez, C., Calvo-Morata, A., Freire, M., Martinez-Ortiz, I., and Fernandez-Majón, B. (2019). Applications of data science to game learning analytics data: a systematic literature review. *Comput. Educ.* 141, 103612–103619. doi: 10.1016/j.compedu.2019.103612

Bertozzi, E. (2014). *Using Games to Teach, Practice and Encourage Interest in STEM Subjects.* Learning, Education and Games. ETC Press, Pittsburgh, PA, USA, 23–36.

Bitner, J. (2021). *What is Minecraft?.* Digital Trends. Available at: https://www.digitaltrends.com/gaming/what-is-minecraft/.

Bonner, D., and Dorneich, M. (2016). Developing game-based learning requirements to increase female middle school students interest in computer science. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 60, No. , pp. 380–384). Sage CA: Los Angeles, CA: SAGE Publications.

Casano, J. D. L., and Rodrigo, M. M. T. (2022a). "A comparative assessment of US and PH learner traversals and in-game observations within Minecraft," in *Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners' and Doctoral Consortium. AIED 2022.* eds. M. M. Rodrigo, N. Matsuda, A. I. Cristea and V. Dimitrova, Lecture Notes in Computer Science, vol. *13356* (Cham: Springer).

Casano, J. D. L., and Rodrigo, M. M. T. (2022b). "An Analysis of Filipino Primary School Learners' Game Experience and STEM Interest within Minecraft," *in 30th International Conference on Computers in Education (ICCE 2022).* eds. S. Iyer, J-L. Shih, W. Chen, and M. N. MD Khambari (Taiwan: Asia-Pacific Society for Computers in Education).

Chu, H., and Chang, S. (2014). Developing an educational computer game for migratory bird identification based on a two-tier test approach. *Educ. Technol. Res. Dev.* 62, 147–161. doi: 10.1007/s11423-013-9323-4

Ekaputra, G., Lim, C., and Eng, K. I. (2013). *Minecraft: A Aame as an Education and Scientific Learning Tool.* Bali, Indonesia: ISICO.

Ennis, L. (2018). Game-Based Learning: An Instructional Tool. Available at: https://dr.lib.iastate.edu/handle/20.500.12876/17144

Esclamado, M. A., and Rodrigo, M. M. T. (2022a). "Are all who wander lost? An exploratory analysis of learner traversals of Minecraft worlds" in *Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners' and Doctoral Consortium. AIED 2022.* eds. M. M. Rodrigo, N. Matsuda, A. I. Cristea and V. Dimitrova, Lecture Notes in Computer Science, vol. *13356* (Cham: Springer)

Esclamado, M. A., and Rodrigo, M. M. T. (2022b). "Achievement, behaviors, and STEM interest of frustrated and bored learners using Minecraft," *in 30th International*

Conference on Computers in Education (ICCE 2022). eds. S. Iyer, J-L. Shih, W. Chen, and M. N. M. D. Khambari (Taiwan: Asia-Pacific Society for Computers in Education).

Gadbury, M., and Lane, H. C. (2022). "Mining for STEM interest behaviors in Minecraft," in *Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners' and Doctoral Consortium. AIED 2022*. eds. M. M. Rodrigo, N. Matsuda, A. I. Cristea and V. Dimitrova, Lecture Notes in Computer Science, vol. *13356* (Cham: Springer). doi: 10.1007/978-3-031-11647-6_42

Gao, F., Li, L., and Sun, Y. (2020). A systematic review of mobile game-based learning in STEM education. *Educ. Technol. Res. Dev.* 68, 1791–1827. doi: 10.1007/s11423-020-09787-0

Hussein, M. H., Ow, S. H., Cheong, L. S., Thong, M. -K., and Ale Ebrahim, N. (2019). Effects of digital game-based learning on elementary science learning: a systematic review. *IEEE Access* 7, 62465–62478. doi: 10.1109/ACCESS.2019.2916324

Hwang, G.-J., Sung, H. -Y., Hung, C. -M., Haung, I., and Tsai, C. -C. (2012). Development of a personalized educational computer game based on students' learning styles. *Educ. Technol. Res. Develop.* 60, 623–638. doi: 10.1007/s11423-012-9241-x

IJsselsteijn, W. A., De Kort, Y. A., and Poels, K. (2013). *The Game Experience Questionnaire*. Technische Universiteit Eindhoven.

Iliya, A., and Jabbar, A. (2015). Gameplay engagement and learning in game-based learning: a systematic review. *Rev. Educ. Res.* 85, 740–779. doi: 10.3102/0034654315577210

Ishak, S. A., Din, R., Othman, N., Gabarre, S., and Hasran, U. A. (2022). Rethinking the ideology of using digital games to increase individual interest in STEM. *Sustainability* 14:4519. doi: 10.3390/su14084519

Johnson, D., Gardner, M. J., and Perry, R. (2018). Validation of two game experience scales: the player experience of need satisfaction (PENS) and game experience questionnaire (GEQ). *Int. J. Hum. Comput. Stud.* 118, 38–46. doi: 10.1016/j.ijhcs.2018.05.003

Lane, H. C., Gadbury, M., Ginger, J., Yi, S., Comins, N., Henhapl, J., et al. (2022). Triggering STEM interest with Minecraft in a hybrid summer camp. *Technol. Mind Behav.* 3, 1–18. doi: 10.1037/tmb0000077

Lent, R. W., and Brown, S. D. (2008). Social cognitive career theory and subjective well-being in the context of work. *J. Career Assess.* 16, 6–21. doi: 10.1177/1069072707305769

Leong, P., Eichelberger, A., and Asselstine, S. (2018). Digital building blocks for learning: motivating and engaging students through Minecraft game-based learning. *Int. J. Educ. Media Technol.* 12, 35–41. Available at: https://jaems.jp/contents/icomej/vol12-2/05_Leong_Eichelberger.pdf

Lester, J. C., Spires, H. A., Nietfeld, J. L., Minogue, J., Mott, B. W., and Lobene, E. V. (2014). Designing game-based learning environments for elementary science education: a narrative-centered learning perspective. *Inf. Sci.* 264, 4–18. doi: 10.1016/j.ins.2013.09.005

Manahan, D. M. A. B., and Rodrigo, M. M. T. (2022). "Support structures and activities 1908 for teachers preparing for game-based learning. 10th workshop on technology-enhanced 1909 STEM education (TeSTEM," *in 30th International Conference on 1910 Computers in Education (ICCE 2022)*. eds. S. Iyer, J-L. Shih, W. Chen and M. N. M. D. Khambari (Taiwan: Asia-Pacific Society for Computers in Education).

Mohtar, L. E., Halim, L., Abd Rahman, N., Maat, S. M., Iksan, Z. H., and Osman, K. (2019). A model of interest in stem careers among secondary school students. *J. Balt. Sci. Educ.* 18, 404–416. doi: 10.33225/jbse/19.18.404

Nkadimeng, M., and Ankiewicz, P. (2022). The affordances of Minecraft education as a game-based learning tool for atomic structure in junior high school science education. *J. Sci. Educ. Technol.* 31, 605–620. doi: 10.1007/s10956-022-09981-0

Pusey, M., and Pusey, G. (2015). Using Minecraft in the science classroom. *Int. J. Innov. Sci. Math. Educ.* 23, 22–34. Available at: https://openjournals.library.sydney.edu.au/CAL/article/view/10331

Roller, S. A., Lampley, S. A., Dillihunt, M. L., Benfield, M. P., and Turner, M. W. (2018). "Student attitudes toward STEM: a revised instrument of social cognitive career theory constructs (fundamental)," In *2018 ASEE Annual Conference & Exposition*. (American Society for Engineering Education).

Saricam, U., and Yildirim, M. (2021). The effects of digital game-based STEM activities on Students' interests in STEM fields and scientific creativity: Minecraft case. *Int. J. Technol Educ. Sci.* 5, 166–192. doi: 10.46328/ijtes.136

Shang, J., Ma, S., Hu, R., Pei, L., and Zhang, L. (2019). "Game-based learning in future school" in *Shaping Future Schools with Digital Technology. Perspectives on Rethinking and Reforming Education*. eds. S. Yu, H. Niemi and J. Mason (Singapore: Springer), 125–146.

Wang, L. H., Chen, B., Hwang, G. J., Guan, J. Q., and Wang, Y. Q. (2022). Effects of digital game-based STEM education on students' learning achievement: a meta-analysis. *Int. J. STEM Educ.* 9, 1–13. doi: 10.1186/s40594-022-00344-0

What-If Hypothetical Implementations in Minecraft (WHIMC) (n.d.). What-If Hypothetical Implementations in Minecraft (WHIMC). Available at: https://whimcproject.web.illinois.edu/

Yi, S., Gadbury, M., and Lane, H. C. (2020). "Coding and analyzing scientific observations from middle school students in Minecraft," in *The Interdisciplinarity of the Learning Sciences, 14th International Conference of the Learning Sciences (ICLS) 2020* Vol. 3. eds. M. Gresalfi and I. S. Horn (Virtual Nashville, TN, USA: International Society of the Learning Sciences), 1787–1788.

Yi, S., Gadbury, M., and Lane, H. C. (2021). "Identifying and coding STEM interest triggers in a summer camp," in *Proceedings of the 15th International Conference of the Learning Sciences – ICLS 2021*. eds. E. de Vries, Y. Hod, and J. Ahn (Bochum, Germany: International Society of the Learning Sciences), 915–916.

Yi, S., and Lane, H. C. (2019). *What-if Hypothetical Implementations in Minecraft (WHIMC)*. Available at: https://www.researchgate.net/publication/339997313_What-if_hypothetical_implementations_in_Minecraft_WHIMC

frontiers | Frontiers in Education

TYPE Curriculum, Instruction, and Pedagogy
PUBLISHED 17 March 2023
DOI 10.3389/feduc.2023.1110464

Check for updates

OPEN ACCESS

EDITED BY
Wang-Kin Chiu,
The Hong Kong Polytechnic University,
China

REVIEWED BY
Monique Meier,
Technical University Dresden,
Germany
Grace Ukasoanya,
University of Manitoba,
Canada
Alexandria Hansen,
California State University,
Fresno,
United States

*CORRESPONDENCE
Cíntia Marques Coelho
✉ cintiacoelhom@unb.br

SPECIALTY SECTION
This article was submitted to
STEM Education,
a section of the journal
Frontiers in Education

RECEIVED 28 November 2022
ACCEPTED 03 March 2023
PUBLISHED 17 March 2023

CITATION
Oss Boll H, Leitão MC, Valle Garay A, Batista ACC, Resende SG, Silva LF, Reis VCB, Vieira EM and Coelho CM (2023) SynBio in 3D: The first synthetic genetic circuit as a 3D printed STEM educational resource. *Front. Educ.* 8:1110464. doi: 10.3389/feduc.2023.1110464

SynBio in 3D: The first synthetic genetic circuit as a 3D printed STEM educational resource

Heloísa Oss Boll[1], Matheus de Castro Leitão[1], Aisel Valle Garay[2], Ana Carolina Campos Batista[1], Sophia Garcia de Resende[1], Leonardo Ferreira da Silva[1], Viviane Castelo Branco Reis[3], Emerson Monteiro Vieira[4] and Cíntia Marques Coelho[1]*

[1]Department of Genetics and Morphology, Laboratory of Synthetic Biology, University of Brasília, Institute of Biological Sciences, Brasília, Brazil, [2]Department of Cell Biology, Laboratory of Molecular Biophysics, University of Brasília, Institute of Biological Sciences, Brasília, Brazil, [3]Department of Cell Biology, Laboratory of Molecular Biology, University of Brasília, Institute of Biological Sciences, Brasília, Brazil, [4]Department of Ecology, Laboratory of Vertebrate Ecology, University of Brasília, Institute of Biological Sciences, Brasília, Brazil

Synthetic biology is a new area of science that operates at the intersection of engineering and biology and aims to design and synthesize living organisms and systems to perform new or improved functions. Despite the important role it plays in resolving global issues, instructing synthetic biology can be challenged by a limited availability of specific educational materials and techniques for explaining complex molecular mechanisms. On the other hand, digital fabrication tools, which allow the creation of 3D objects, are increasingly used for educational purposes, and several computational structures of molecular components commonly used in synthetic biology processes are deposited in open databases. Therefore, we hypothesized that the use of computer-assisted design (CAD) and 3D printing to create biomolecular structural models through hands-on interaction, followed by reflective observation, critical and analytical thinking, could enhance students' learning in synthetic biology. In this sense, the present work describes the design, 3D printing process, and evaluation in classrooms of the molecular models of the first synthetic biological circuit, the genetic toggle switch. The 3D printed molecular structures can be freely downloaded and used by teachers to facilitate the training of STEM students in synthetic biology. Most importantly, the results demonstrated that our resource showed a significant positive impact ($p<0.05$) on students' learning process, indicating that the proposed method helped them better understand the genetic toggle switch.

KEYWORDS
synthetic biology (synbio), genetic toggle switch, 3D printing, stem education, educational resource

Introduction: Background and rationale for the educational activity innovation

An overview on synthetic biology and the first synthetic genetic circuit

Synthetic biology comprises a series of disruptive technologies capable of providing new solutions to global challenges in health, agriculture, industry, and the environment (Cameron et al., 2014; Flores Bueso and Tangney, 2017; French, 2019). In its top-down approaches, such an area uses molecular biology tools and techniques coupled with standardized engineering principles to design and characterize biological parts, synthetic biological circuits, and systems. Thus, synthetic biology can assign new functions to organisms and redesign pre-existing biological systems to improve features of interest (Fu, 2006; Huang et al., 2010; Lv et al., 2022).

Jacob and Monod described, in the 1960s, the first genetic circuit, as they observed that inducible systems, such as the lactose operon from *Escherichia coli*, and repressible systems, like the tryptophan operon or lysogenic systems, respond to similar controlling elements, organized in different genetic circuits (Jacob and Monod, 1961). For example, in the lactose operon, the authors observed that the lactose catabolism was triggered only by a specific inducer molecule, structurally similar to lactose, which stimulated the coordinated synthesis of enzymes that allow lactose processing. On the other hand, when lactose is absent and glucose is present in the medium, the last sugar is preferred, due to easier processing, and the lac operon is silenced by the association of a repressor protein with the operator region of the lac promoter. The description of inducer and repressor components that could genetically control cellular machineries subsequently enabled, decades later, the development of synthetic biological circuits.

The progress made in the field of genetic engineering was also instrumental for the development of synthetic biology. In the 1970s, the discovery of restriction enzymes and the development of recombinant DNA technology allowed for the controlled manipulation and expression of DNA fragments in living organisms (Luria, 1970; Smith and Wilcox, 1970). This equipped researchers with the means to better regulate gene expression, leading to a deeper understanding of gene function. The foundations of synthetic biology are therefore closely intertwined with genetic engineering, as the latter has provided the fundamental molecular principles and DNA manipulation techniques needed to implement the former. The capability of synthetic biology to modify numerous nucleotides or gene *loci* across the genome sets it apart from conventional genetic engineering, which is limited to altering a limited number of nucleotides or genes typically using recombinant DNA technology (König et al., 2013). In the early 2000s, initial attempts were made to expand the frontiers of genetic engineering and create synthetic genetic circuits.

The first synthetic genetic switch assembled was based on two genes that repressed each other (Gardner et al., 2000). The application of a specific stimulus induced the expression of one of the genes, and its genetic product inhibited the expression of the other, and vice versa. One of the genes was placed *in tandem* with a reporter sequence encoding the green fluorescent protein (GFP), and thus the two possible states of the system were the presence or absence of green fluorescence. Due to its similarity to an electronic toggle switch, which also has two possible states—"ON" (light) or "OFF" (no light)—their system is recognized as the genetic toggle switch. Shortly after, Elowitz and Leibler developed the genetic oscillator, a circuit characterized by the interactions among three different genetic repressors, one of which also regulated GFP expression (Elowitz and Leibler, 2000). At a cellular level, the circuit caused periodic oscillations in the detection of green fluorescence, similar to what happens in electronic oscillators. Both studies strengthened the foundations of synthetic biology and demonstrated that it is possible to apply engineering principles to create synthetic biological circuits in living organisms.

Current challenges in synthetic biology education

Teaching synthetic biology presents unique challenges compared to other engineering fields. Electrical engineering students, for example, can easily obtain parts of electrical circuits, such as conductor wires, electronic switches, and LEDs, to learn in practice the concepts studied during theoretical classes—such as building a simple electronic toggle switch to turn on a light bulb. In contrast, synthetic biology focuses on designing and building biological circuits, which operate at the molecular scale. Consequently, it is harder for students to connect the concepts learned in class with both the actual dimension of subcellular processes and the relations between biological genetic processes and engineering. One of the main reasons for that is the scarcity of specific pedagogical methods for teaching synthetic biology (Diep et al., 2021).

To date, efforts have been made to incorporate novel resources into synthetic biology education, such as software for molecular structure simulation and visualization, virtual labs, and virtual gaming environments (Muth et al., 2021). For example, the Pymol software helps in understanding protein shapes and their relation to specific functions, and is used by molecular science instructors for visualization and computation of structures (Lineback and Jansma, 2019). Virtual labs, such as Serial Cloner, provide a whole DNA assembly virtual environment, while GelBox offers an interactive simulation tool for gel electrophoresis, both of which are essential techniques in genetic engineering and synthetic biology (Basics, 2009; Gingold and Douglas, 2018). Virtual gaming environments, such as Hero.coli, which reproduces a bacteria incorporating genetic elements to gain new properties, and Nanocrafter, which simulates the assembly of DNA fragments to create new genetic devices, help students learn basic concepts of genetic transformations in a fun manner (Barone et al., 2015; Goujet, 2018). Interactive experimental methods, like BioBits, which allows students hands-on exposure to synthetic biology experiments involving fluorescence, fragrances, and hydrogels, have also been proposed for synthetic biology education (Huang et al., 2018, 2022).

Despite these efforts, the teaching of synthetic biology still faces challenges similar to those faced in the teaching of genetic engineering and molecular sciences. One of the main difficulties lies in transforming static, two-dimensional illustrations present in scientific articles and textbooks into dynamic, three-dimensional models that truly bring the subject to life (Wu et al., 2001). Unfortunately, the fact that many students have difficulty with three-dimensional mental visualizations is often overlooked and can result in a disadvantage in their careers in STEM fields (Pittalis and Christou, 2010). In this sense, the previously mentioned educational resources still have limitations in fully demonstrating molecular processes in three dimensions, which is crucial for students' understanding of synthetic biology.

An overview of 3D printed molecules and their potential in education

In the early 1950s, Linus Pauling, Robert Corey, and Herman Branson, pioneers in studies of protein structures, developed the first representation of macromolecules in terms of complex space-filling models (Pauling et al., 1951; Pauling and Corey, 1951). Their system represented atoms of different chemical elements as spheres of different colors, whose diameter was proportional to their atomic radius (Koltun, 1965; Olson, 2018). In 1958, researchers reported the first experimental structure of a macromolecule, and for 20 years physical models were the principal tool for representing the structures of biological macromolecules (Kendrew et al., 1958; Olson, 2018).

Physical models fell into disuse in the 1980s, when visualization of biomolecular structures by molecular computer graphics softwares became popular. Fortunately, in recent years, several independent efforts have created physical models of proteins using new rapid prototyping technologies based on 3D printing. 3D printers enable the manufacture of macromolecule structures using digital atomic coordinates, encoded in *.pdb* files and available for download from the Protein Data Bank (PDB), with low waste and high accuracy and efficiency, simplifying the process of making molecular models.

Recently, about 47 peer-reviewed articles were systematically reviewed to identify different justifications for incorporating 3D printing into higher education chemistry (Pernaa and Wiedmer, 2019). These justifications include addressing challenges in chemistry learning and teaching, overcoming the high cost and lack of suitable molecular models, and the limitations of current molecular models. The majority of the articles (about 37) focused on specific chemistry concepts or laboratory instruments that could benefit from 3D printing, as well as the development of printing methodologies, safety considerations, and pedagogical models to evaluate the impact of physical models on student learning and perception.

In addition, tactile feedback has already been found to be more valuable than 3D representations alone for students struggling to understand molecular biological concepts (Salzman et al., 1999). 3D printing has been explored in recent reviews for its potential in science education, with studies reporting increased biological and chemical conceptual gains for students using 3D printed models (Pinger et al., 2019; Hansen et al., 2020). The use of three-dimensional educational materials is particularly important for visually impaired students who face substantial barriers in the classroom due to a lack of tactile methods, which are fundamental for better understanding biological concepts (Stone et al., 2020).

Study hypothesis

We hypothesize that the use of 3D printed biological structures that comprise the genetic toggle switch can enhance students' learning of synthetic biology. By allowing them to manipulate and understand theoretical concepts in a practical way, the 3D printed parts can provide a comprehensive educational experience that captures student interest and clarifies complex molecular concepts. By using this open resource, professors would be able to show biological parts on an adequate scale, helping students better understand macromolecule spatial relationships and genetic mechanisms and their relations to electrical engineering in a way that illustrations and computational models alone cannot.

Development of the 3D printed educational resource

To construct our educational resource, we selected the first genetic toggle switch, which was designed, synthesized, and successfully tested by Gardner et al. (2000). First, we identified the molecular elements that compose the switch: two molecular inducers, anhydrotetracycline (aTc) and isopropyl β-D-1-thiogalactopyranoside (IPTG); two repressor gene products, the lactose operon repressor (LacI) and the Tet repressor protein (TetR); and a reporter gene product, the green fluorescent protein (GFP). As the system depends on gene expression, we also included an RNA polymerase, to allow the understanding of the control of transcriptional activity, and two generic DNA strands, each to represent the two mutually repressive genetic cassettes.

Figure 1 shows a schematic representation of the steps to follow to obtain the 3D printed molecules of the genetic toggle switch. The *.stl* files encoding the 3D digital atomic coordinates of the proteins were downloaded from the Protein Data Bank (PDB; http://www.pdb.org/pdb/). For the smaller molecules, the. M*ol* files were downloaded from the Molview database. Structures were prepared in the PyMol and ChimeraX v1.3 (https://pymol.org/2/; https://www.rbvi.ucsf.edu/chimerax) software by extracting the asymmetric units from the unitary cells, removing waters and ligands, and calculating the molecular surfaces. Afterwards, the molecular surface structures were exported as *.stl* files using ChimeraX, or as VRML 2 files from PyMol for repair, and hollowed out with Autodesk Meshmixer v.3.5.474 or Blender software. In Blender, each element was imported individually as X3D Extensible 3D files. The skeleton of each piece was deleted, preserving only the empty surface shells. In Meshmixer, the *.stl* files of the molecular surfaces were hollowed out and repaired to generate the final *.stl* files, for use in the UltimakerCura slicer.

For slicing in the UltimakerCura software, we scaled the *.stl* molecular files into two size scales. The inducers were printed at 100% of their initial size, and the other components at 70% of scale. These scales were chosen to facilitate manipulation of the biological structures by students while maintaining approximate natural proportions. The 3D prints were made on a fused deposition (FDM) 3D printer with a 0.4 mm diameter extrusion nozzle using 1.75 mm polylactic acid plastic (PLA) filaments. In the slicing software, we set the following parameters: (1) the extrusion width was set according to the extrusion nozzle diameter; (2) the printing temperatures ranged from 200 to 215°C, (3) the printing speed ranged from 20 to 60 mm/s, and the temperature and speed parameters varied according to the individual size of each printed part. 200 g of PLA filaments were used in yellow (LacI), white (IPTG), blue (TetR), light blue (aTc), gray (DNA), green (GFP), and pink (RNA polymerase). After printing, printing supports were removed from the models by hand, and small 5 × 5 mm neodymium magnets were added approximately at the real sites of the biological interactions to allow reversible binding between the pieces representing the genetic toggle switch molecules and better represent the activation and repression behaviors of transcription.

Pedagogical framework, learning environment, objectives, and core competencies development

Pedagogical background

The pedagogical concept of this study is based on Kolb's Experiential Learning Theory (ELT), which has been discussed as being built upon the works of Dewey, Lewin, and Piaget (Kolb, 1984).

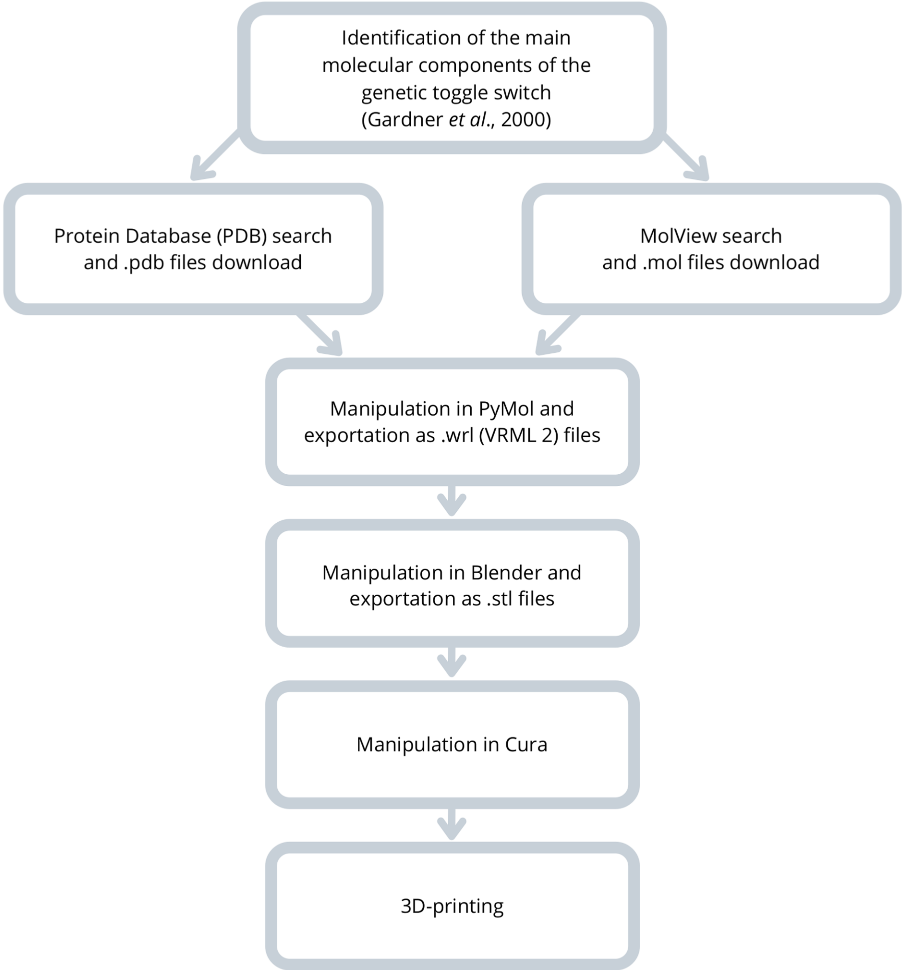

FIGURE 1
Schematic representation of the pipeline used for production of the genetic toggle switch as a system of 3D printed pieces.

The ELT combines four learning styles into a four-stage learning cycle comprising (1) concrete learning, where students encounter a new experience; (2) reflective observation, where students reflect on the new experience while considering their existing knowledge; (3) abstract conceptualization, where students give rise to new ideas based on the previous considerations; and (4) active experimentation, where students experiment and apply the newly created knowledge in real situations. It is argued that students can achieve a better understanding of concepts through the challenge present in problem solving, reinforcing and enhancing learning and critical thinking; therefore, it underpins the chosen teaching method described here (Kolb, 1984; Kolb et al., 2001).

The approach proposed here has already been tested by other authors in the broader field of STEM education, in which synthetic biology is included. A number of significant findings highlight that active engagement learning strategies have already been shown to reduce the percentage of failure rates compared to regular lectures in undergraduate courses in STEM subjects, and are also associated with a statistically significant improvement in individual learning performance and an increase in average assessment scores in molecular science (Newman et al., 2018). Specifically in the context of molecular biology, the use of tactile models has been shown to increase learning gains related to the central dogma of molecular biology, DNA replication and transcription, and protein folding, for example (Beltramini et al., 2006; Davenport et al., 2017; Gordy et al., 2020).

Participants and learning environment

Our approach focused on assessing whether 3D printed molecules could enhance students' understanding of synthetic biology, specifically around the first synthetic biological circuit, the genetic toggle switch. Thus, the learning environment was the classrooms of the University of Brasília, and our participants were undergraduate students majoring in Biotechnology and Biological Sciences. More specifically, the participants were students who were already at least halfway through their respective courses and who were enrolled in one of the Genetic Engineering, Genetics or one of the two Molecular Biology classes (A and B) offered that semester. Besides that, the only prerequisite for participation was that the students had to have attended a class in which they studied the lactose operon, since the genetic regulation of this operon was the foundation for the development of the first synthetic genetic circuit. The number of students in each of the four classes evaluated was 21, 23, 8, and 33 respectively, resulting in a total of 85 students.

Figure 2 describes a pipeline of the learning environment, the steps involving the application of the Pretests and Posttests, and the relations of the educational resource proposed in this study to the ELT, which underlies the chosen teaching method.

Examination procedure

Our main objective was to develop a new educational resource for synthetic biology that would improve student learning. To accomplish this goal, the coordinating professors of the study, with the assistance of undergraduate and graduate students serving as teaching assistants, consistently implemented the same teaching strategy and procedures across all tested classes. All classes lasted for 2h. During the first 45min, the professors in charge explained that the study was being conducted according to the ethical standards set by the University of Brasília, and they were assured that their participation would be private, confidential, and voluntary. They were also informed that their identity would remain anonymous and that the data collected would be used exclusively for academic purposes, and treated with the utmost confidentiality. The participants signed an agreement form acknowledging their participation and acceptance of the conditions outlined in the research methodological process. Next, the professors in charge, using the traditional lecture method based on the whiteboard and slide decks, introduced all the molecules and genetic parts involved in the genetic toggle switch, and explained how they operated. Participants were randomly numbered and asked to answer a Pretest questionnaire using the Pretest-Posttest questionnaire (Supplementary material). They were given 15min to perform this task, synchronously. The questionnaire was composed of multiple choice and true and false questions.

Afterward, the participants were divided into two groups (control and experimental). The second part of the class lasted 45min for both groups (Posttest). The students who were randomly assigned as even

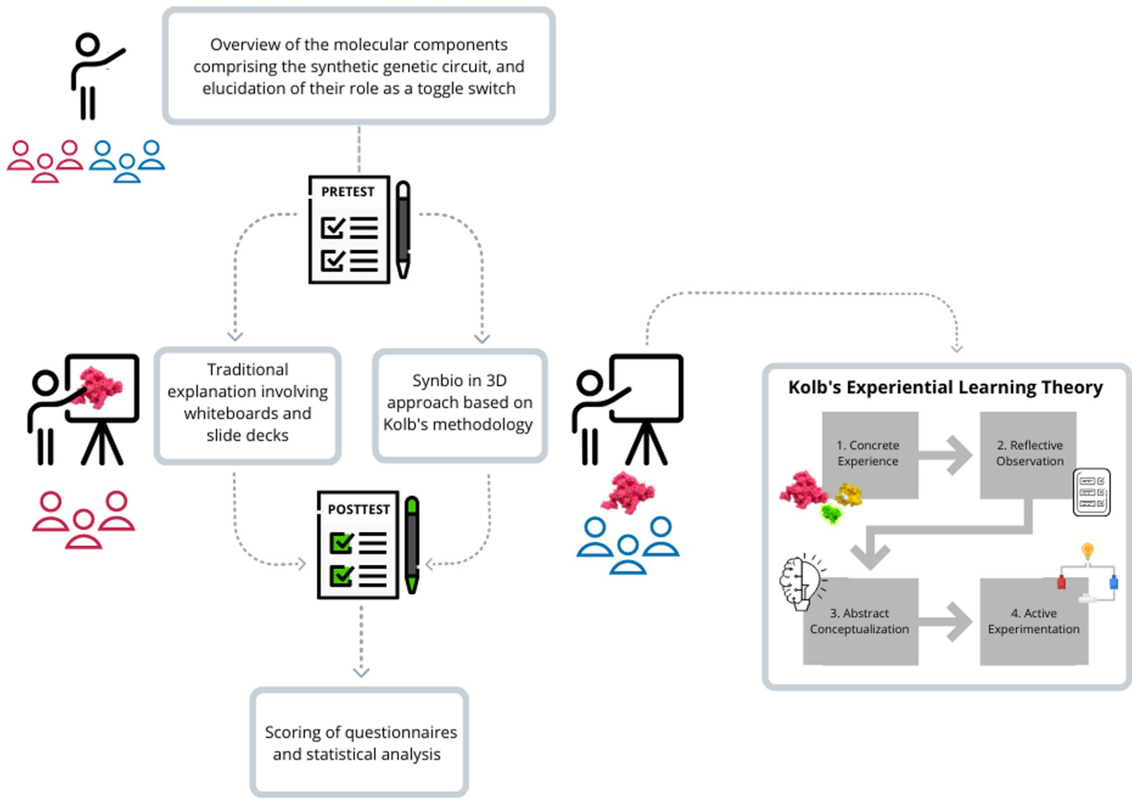

FIGURE 2
Pipeline describing the learning environment. First, the responsible professors made clear that the study was conducted according to the ethical standards set by the University of Brasilia, Brazil. Then, students were randomly numbered and this first part of the class lasted 45min (pretest). Using the traditional lecture method, the molecular components that compose the toggle switch were introduced, followed by an explanation of how these components control gene expression of two repressors proteins and the fluorescent reporter protein setting the ON and OFF cellular state. The students responded to the Pretest-Posttest questionnaire at the last 15min of the first part of the class. In the second part of the class (posttest), which also lasted 45min, the participants were divided into two groups. One group (control group) received the same traditional explanation and responded to the Pretest-Posttest Questionnaire in 15min. The other group (experimental group) received the 3D printed biological structures that compose the toggle switch and were asked to manipulate the models to understand and explain to each other how the toggle switch functions to set the ON and OFF cellular state according to Kolb's Experiential Learning Theory. The control group and experimental group students are represented by red and blue colors, respectively, in the figure. They responded to the Pretest-Posttest Questionnaire in 15min. The data were analyzed, right answers scored 1 and wrong answers 0, and submitted to a non-parametric Mann–Whitney test.

numbers composed the control group, and were asked to move to a different classroom and were given the same traditional explanation. They were then asked to respond to the Pretest-Posttest questionnaire, also in 15 min.

For the remaining students, who were assigned as odd numbers and who comprised the experimental group, the professors in charge reproduced the following steps in front of the class with the 3D printed molecules:

1. Introduce again all the molecules and genetic parts involved in the synthetic genetic circuit, describing that IPTG and aTc are the inducers and LacI and TetR are the repressor proteins for pTrc-2 and $P_L$tetO-1, respectively. As promoters, $P_L$tetO-1 and pTrc-2 control the expression of LacI and TetR together with GFP, respectively, in the two different expression cassettes.
2. Simulate how RNA polymerase transcribes both of the cassettes. In the first case, the magnetic connection of the RNA polymerase with the pTrc-2 promoter region of a DNA strand causes the expression of TetR and GFP. In the second case, the transcription leads to the expression of LacI. These final products were not be visible to the class until transcription has been simulated. However, the translation process omitted in our system was highlighted as the step responsible for converting the information contained in the expressed mRNAs into individual proteins.
3. Show how the addition of the two repressors, TetR and LacI, blocks RNA polymerase activity at each of the two promoters in the two expression cassettes. Repeat the last step, but explain that a physical attachment of LacI or TetR on the promoter region of pTrc-2 or $P_L$tetO-1, respectively, impedes the transcription of each cassette by RNA polymerase. No transcription and consequently no translation takes place, so the class did not get to see the final product. At his point, the responsible professors stated that it is important to note that in the system, the binding of only one molecular repressor per promoter is a simplified representation of the cell's reality; many more molecules operate to block promoter regions.
4. Explain how the two cassettes interact with each other.
   a. Draw a parallel with a standard light switch that can set a simple electrical system into two possible states: the presence (ON) or absence (OFF) of light. Thus, the genetic switch can set the biological system into two possible states: the presence (ON) or absence (OFF) of green fluorescence. In this case, the responsible professors stated that it is worth noting that GFP is the product that confers this characteristic.
   b. Draw a parallel between the two cassettes and show that IPTG, once added to the medium, is present inside the cell, and that the repressor product of the second cassette, LacI, has the property of binding to this molecule, using the magnets embedded in both parts. Then, the responsible professores asked the students: if LacI is not bonded to pTrc-2 due to the interaction with IPTG, what happens to the expression of the pTrc-2 cassette? As a consequence, the transcription process was shown again, which results in the synthesis of TetR and GFP.
   c. Highlight the TetR produced by the first cassette. Explain that this second molecular repressor attaches to the $P_L$tetO-1 promoter of the second cassette and blocks its expression. Consequently, LacI is no longer produced. Finally, the teacher asked: "What state does the system assume when IPTG is added to the medium?" The final answer was: "The system is in its ON state and shows a green fluorescence."
   d. Explain how the second state is reached by showing the property of the molecular repressor aTc to bind to TetR by a magnetic connection. Then the teacher asked the students: "If, after its addition to the medium, aTc interacts with TetR, what happens to the expression of the $P_L$tetO-1 cassette?" In this case, the transcription process is restarted, leading to the synthesis of LacI, a repressor of the pTrc-2 promoter. When the pTrc-2 is blocked, there is no expression of GFP and TetR. Finally, the teacher asked: "What state does the system assume when aTc is present in the medium?" The final answer was: "The system is OFF and colorless."

Then, the students in the experimental group were divided into small groups of two, given the 3D models, and asked to individually reproduce how biological structures interact to regulate gene expression, leading to ON and OFF states. Following this step, students were asked to reflect and reproduce to each other what they had learned, proceeding with reflective observation, analytical and critical thinking, and verifying if the concepts were indeed fully understood. Students were also encouraged to correlate what they had learned with electrical circuits, encouraging them to expand their knowledge to more complex circuits. Immediately after, they were asked to respond to the Pretest-Posttest questionnaire, also in 15 min.

Evaluation approach

The questionnaires were scored as follows: correct answers were graded with one point, and wrong answers with zero points. After scoring the Pretest and Posttest questionnaires for the four classes in both control and experimental conditions, we conducted a difficulty analysis. The difficulty scores for the Pretest and Posttest were calculated by dividing the number of correct answers by the total number of questionnaires answered in each class and condition.

Learning objectives and expected core competencies development

The learning objective of the class was to enhance the students' comprehension of the control of gene expression underlying the two different states of the toggle switch, ON and OFF, through manipulation of the 3D printed biological structures that compose the genetic circuit, and reflective observation, analytical and critical thinking. At the end, it is expected that the participants can apply the new ideas learned from the proposed method to understand more complex biological circuits, and in the future use this knowledge to develop projects in the realm of synthetic biology.

The proposed educational activity aimed to foster the STEM competencies expected for the 21st century, including problem-solving, critical thinking, creativity, innovation, communication, and collaboration, as identified by the UNESCO report on STEM competencies for the 21st century (UNESCO International Bureau of Education, 2019). The SynBio in 3D activity relies on the manipulation of 3D printed biological structures by students, mimicking the inductors and protein functions inside the cell. Through explaining the structures to each other in a reflective and analytical way, it is hypothesized that the students will better understand the genetic control underlying the toggle switch, and generate abstract principles that can be applied in more complex circuits. Therefore, the competencies underlying this learning activity encompass both hard and soft skills. Specifically, the core competencies would be the ability and willingness to learn, conceptual/critical thinking, teamwork and cooperation, analytical thinking, digital literacy, cultural awareness, and social responsibility, all of which are essential for preparing undergraduate students for the complex and rapidly changing STEM landscape of the future. By drawing upon the UNESCO report on STEM competencies for the 21st century, the educational activity was designed to effectively promote these competencies, equipping students with the skills and knowledge they need to succeed in their studies.

Results to date/assessment (processes and tools; data planned; or already gathered)

3D printing of the educational resource

The rationale for this proposal relies on the hypothesis that students would benefit from 3D models to better understand concepts of genetic circuits. We selected the toggle switch as a starting point, not only because it was the first synthetic circuit assembled, which paved the way for further designs, but also due to its simplicity: it involves only a negative control that relies on two repressors to regulate the expression of the reporter GFP gene. Figure 3 describes all genetic parts that comprise this switch, comparing the computational models experimentally obtained and submitted online and the 3D printed pieces. By preserving the dimensions of each molecule, 3D printed molecular inducers (IPTG and aTc) were significantly smaller than 3D printed protein repressors (TetR and LacI). Also, we used filaments colored in similar tones to the products derived from each of the cassettes, so students could easily associate which system is which.

After preparation, the final genetic circuit parts were 3D printed and finalized with magnets. All files are available in the Supplementary material, on our laboratory's website, synbiolabunb.com, and under the CC BY 4.0 license. This authorizes the work to be shared and adapted under conditions of giving credit to the original authors, not using the materials for commercial purposes, and not applying legal terms or technological measures that legally restrict others from doing anything the license permits.

Molecular mechanism of the genetic toggle switch

Although Jacob and Monod had pointed out in the '60s the existence of biological regulatory circuits, it was not until the beginning of the new millennium that Gardner and collaborators reported the first design and test of a synthetic genetic circuit, the toggle switch (Jacob and Monod, 1961; Gardner et al., 2000).

Figure 4 describes the different cellular states of the genetic toggle switch dependent on its intracellular inducer content. It shows how the genetic control expression is linked to Boolean algebra function as binary variables, considering 1 when the gene is expressed and 0 when it is not. First, it explains the design of the synthetic genetic circuit to control the expression and accumulation of the reporter protein, GFP. With the IPTG addition to the medium, it diffuses within the cells and binds to the LacI protein. This way, the operator of the pTrc-2 promoter is free, allowing the flow of the RNA polymerase—a process called a transcription "current," analogous to an electric current as a transistor-like device, named transcriptor. The result is the expression of the cassette composed of TetR and GFP. TetR binds to the promoter of the other cassette, $P_L$tetO-1, blocking the expression of LacI, and GFP sets the "ON" state of the system, making the cells turn fluorescent green. Then, the figure illustrates how the cells can change state and turn colorless. When aTc is added to the medium, it enters the cell and binds to the TetR protein, leaving the operator of the $P_L$tetO-1 promoter available for the RNA polymerase to transcribe. Hence, LacI is produced, and it blocks the promoter of the other cassette, pTrc-2, setting the "OFF" state of the system: that is, cells are unable to express GFP and turn colorless. In summary: IPTG causes the cells to turn fluorescent green, whereas aTc has the opposite effect, maintaining them colorless.

Classroom evaluation data analysis

The biological molecules that compose the genetic toggle switch were 3D printed and presented to undergraduate students enrolled in Genetics, Genetic Engineering, or one of the two Molecular Biology classes offered at the University of Brasilia that semester. As explained in the previous section, to determine whether the use of 3D printed molecules improved students' comprehension of the genetic toggle switch, they were first given the Pretest-Posttest questionnaire (Supplementary material) after receiving the traditional theoretical explanation (using only a whiteboard and slide decks). This phase was referred to as the Pretest. Then, the students were separated into two groups (Posttest): one group was given the traditional explanation again (control group), while the other group reviewed the concepts with the aid of 3D printed molecules (experimental group).

After scoring the Pretest and Posttest questionnaires (Table 1), a difficulty analysis (Table 2) was conducted. As all groups had a difficulty score above 50%, the test was found to have an intermediate or moderate difficulty, allowing us to proceed with data analysis. It is important to note that in most classes, the experimental conditions had higher difficulty scores than the control conditions. Since difficulty score values are inversely related to test difficulty, these findings suggested that the use of 3D printed models was a valuable educational tool that improved students' comprehension of the genetic toggle switch.

Nonetheless, given the Pretest-Posttest study design employed in this investigation, we sought to examine whether differences in learning existed between the control and experimental groups. To achieve this, we calculated the difference between the Pretest and Posttest scores for each person, and then analyzed these differences (i.e., gain score approach, *sensu* Gliner et al., 2003). To compare these

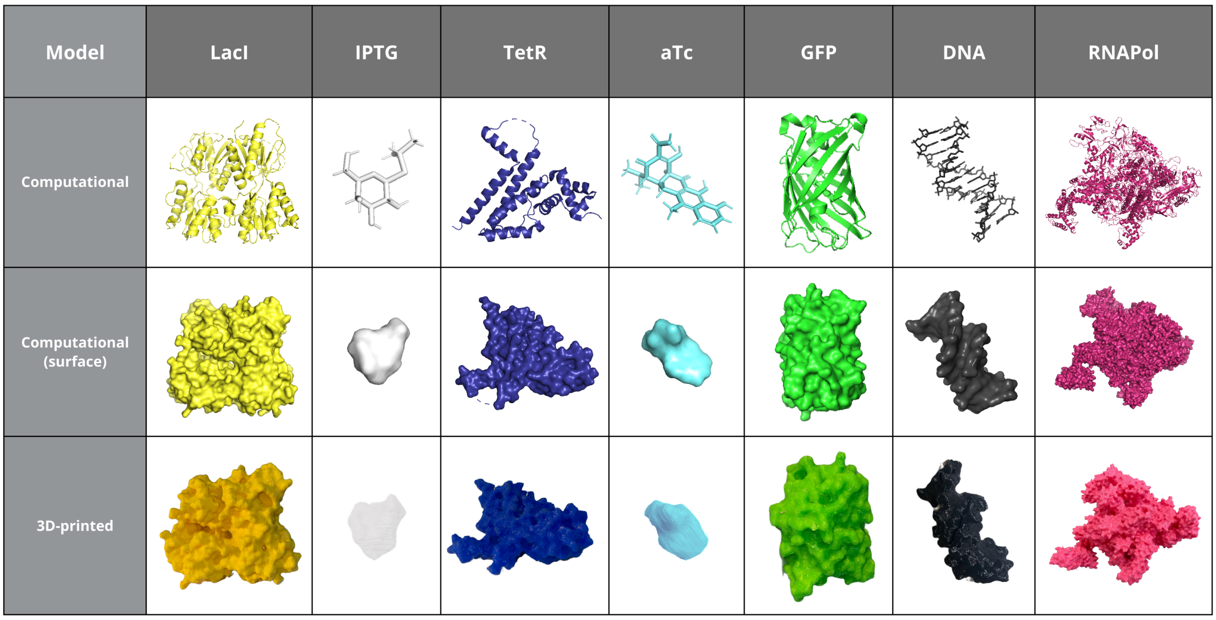

FIGURE 3
Visual comparison of 3D printed models that compose the genetic toggle switch in relation to their respective computational structures.

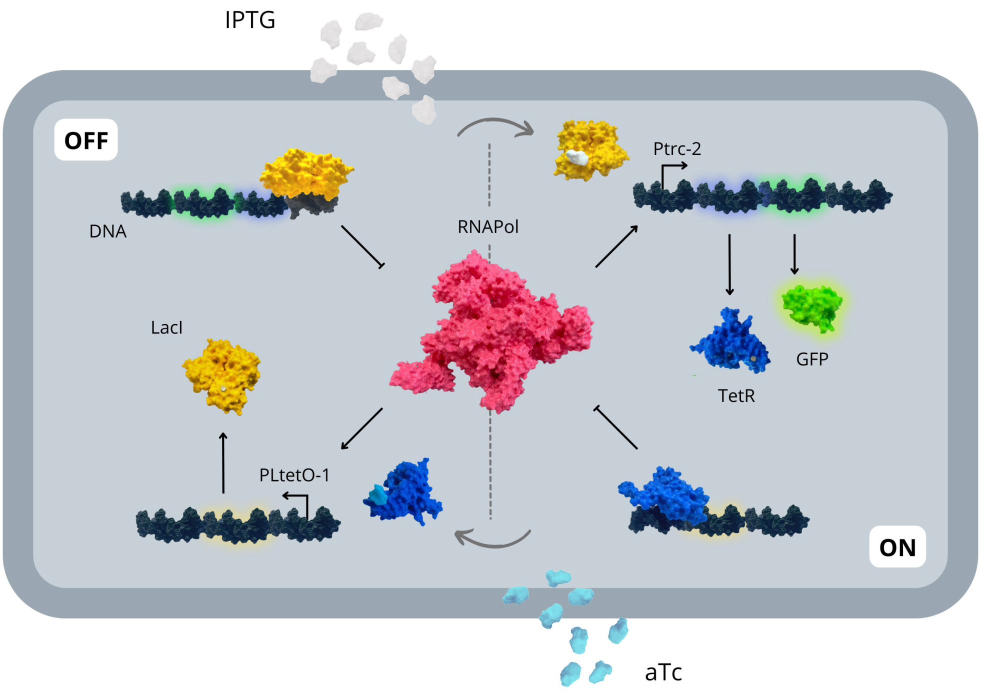

FIGURE 4
Different cellular states that compose the toggle genetic circuit in 3D printed molecules. The molecules and the genetic parts needed in the synthetic genetic circuit to turn it on its ON state are: IPTG (the inducer), LacI (repressor protein), and pTrc-2 (promoter). In its initial state, the LacI repressor is bound to the pTrc-2 promoter, blocking the RNA polymerase flow. After adding IPTG to the medium, it diffuses into the cellular environment, binds the LacI protein, releasing it from the pTrc-2 promoter, and allowing the function of the transcriptor. The cell is now in its ON state, expressing GFP and TetR. The TetR repressor protein binds to the $P_L$tetO-1 promoter and blocks the RNA polymerase flow along the cassette. Then, now, to switch to the OFF state the molecules and the genetic parts needed will be aTc (inducer), TetR (repressor protein), and $P_L$tetO-1 (promoter). Once the aTc is added to the medium, it diffuses into the cellular environment, binds the TetR protein, releasing it from the $P_L$tetO-1 promoter, and allowing the function of the transcriptor. The cell is now in its OFF state, expressing LacI and thus repressing the first cassette from transcribing and traducing TetR and GFP.

TABLE 1 Mean and standard deviation of the students' Pretest and Posttest scores (average number of correct questions) on the four different classes evaluated in this study.

Courses condition	Conditions	Pretest Mean±SD	Posttest Mean±SD	Mean difference
Genetic Engineering	Treatment Control	7.50 ± 2.01	9.78 ± 1.31	2.28
		7.07 ± 1.94	8.33 ± 2.64	1.26
Molecular Biology (A)	Treatment Control	7.75 ± 1.42	10.58 ± 1.68	2.83
		8.64 ± 1.57	10.73 ± 2.20	2.09
Molecular Biology (B)	Treatment Control	6.45 ± 2.07	7.00 ± 2.57	0.55
		7.00 ± 2.49	7.50 ± 3.06	0.50
Genetics	Treatment Control	6.50 ± 1.73	10.75 ± 0.50	4.25
		7.67 ± 2.52	8.33 ± 1.53	0.66
Overall	**Treatment Control**	**7.22 ± 1.88**	**9.40 ± 2.22**	**2.18**
		7.54 ± 2.09	**8.79 ± 2.79**	**1.25**

differences, we employed a non-parametric test (Mann–Whitney test) since the data distribution was non-normal (Shapiro–Wilk test $W = 0.97$; $p < 0.05$). The gain score analysis is commonly considered a good approach for Pretest-Posttest comparisons, considering that there are no Pretest differences between groups (Gliner et al., 2003; Zientek et al., 2016).

The evaluation of the Pretest scores revealed no significant differences between students in the control and experimental groups (Mann–Whitney $U = 769$, d.f. = 82, $p = 0.33$). As a result, we were able to use gain scores to assess the efficacy of the treatment by evaluating between-groups differences. The test indicated significant differences between the two groups, with students in the experimental group achieving a median gain score that was 40% higher than that of students in the control group (Mann–Whitney $U = 643$, d.f. = , $p = 0.03$; Figure 5).

In addition, the qualitative perception of students' comprehension of the toggle switch was evaluated after they attended the lecture (Supplementary Figure 1). The results showed that in the experimental group, after manipulating the 3D printed molecules (phase 2), a higher number of students perceived an improvement in their understanding of the genetic toggle switch compared to the control group in both phases 1 and 2. Moreover, the subjective response to question 8 (optional) of the Pretest-Posttest questionnaire also corroborates the latter results, as most students in the experimental group reported that their comprehension of the genetic toggle switch and the gene regulation control required for its function improved by using 3D printed molecules as an educational resource in the classroom (Supplementary Figure 2).

Discussion on the practical implications, objectives, and lessons learned

Genetic circuits are a combinatorial network of switches that can perceive different inputs, process the information, and generate outputs. Since the beginning of the 2000s, several simple synthetic circuits have been designed, built, and tested (Elowitz and Leibler, 2000; Gardner et al., 2000). These were the cornerstone for assembling more complex ones, which led to further implications in biotechnology, medicine, agriculture, and food sectors, bringing the synthetic biology area to a pivotal importance in the 21st century (Khalil and Collins, 2010; El Karoui et al., 2019; Brooks and Alper, 2021). Therefore, it is of great interest that future researchers in the different areas of science have a thorough understanding of genetic circuits.

The traditional educational system still fosters a passive learning process in which students are stimulated to hear explanations, take notes and then take exams to prove their knowledge. However, we hypothesized that complex and emerging areas, such as synthetic biology, cannot be fully understood without a modern and practical educational approach. The process of elaborating a synthetic biology project is deeply amalgamated into digital technologies, such as apps made to design plasmids from scratch and machines that automate strain engineering. Hence, classes dedicated to synthetic biology education also need to adapt and integrate the use of technology, following up with the tendency and allowing students to usufruct from active learning methodologies.

Indeed, progress in teaching synthetic biology has resulted in the incorporation of new resources, such as virtual laboratories, software for simulating molecular structures, and interactive experimental methods (Huang et al., 2018; Lineback and Jansma, 2019; Muth et al., 2021). Additionally, 3D printed molecular structures are now shown to be also a valuable tool in improving synthetic biology education. The novel SynBio in 3D method was successfully implemented in classrooms, resulting in a positive impact on students' learning processes (Figure 5). It helped students understand the genetic toggle switch and the relationship between gene expression regulation control and electrical circuits, turning cells "ON" and "OFF." These results are consistent with the students' perception that manipulating the 3D printed molecules helped improve their understanding of the subject (Supplementary Figures 1, 2). Thus, our work demonstrates the success of integrating a new type of technology into synthetic biology education by linking digital fabrication tools and molecular computational structures to create engaging content. This innovative approach has the potential to transform traditional educational practices, offering students a modern and interactive learning experience that enhances understanding of emerging areas such as synthetic biology.

TABLE 2 Difficulty analysis of the questionnaires used to evaluate the four different classes.

Course (Class)	Test type	Correct answers	Total	Difficulty score
Genetics	Pretest (control)	2	4.00	0.5
Genetics	Posttest (control)	2	4.00	0.5
Genetics	Pretest (experimental)	2	4.00	0.5
Genetics	Posttest (experimental)	3	4.00	0.8
Genetic Engineering	Pretest (control)	9	15.00	0.6
Genetic Engineering	Posttest (control)	9	15.00	0.6
Genetic Engineering	Pretest (experimental)	11	18.00	0.6
Genetic Engineering	Posttest (experimental)	13	18.00	0.7
Molecular Biology (A)	Pretest (control)	5	10.00	0.5
Molecular Biology (A)	Posttest (control)	6	10.00	0.6
Molecular Biology (A)	Pretest (experimental)	5	11.00	0.5
Molecular Biology (A)	Posttest (experimental)	7	11.00	0.6
Molecular Biology (B)	Pretest (control)	7	11.00	0.6
Molecular Biology (B)	Posttest (control)	8	11.00	0.7
Molecular Biology (B)	Pretest (experimental)	7	12.00	0.6
Molecular Biology (B)	Posttest (experimental)	10	12.00	0.8
-	Total average–Pretest	5.91	10.63	0.56
-	Total average–Posttest	7.47	10.63	0.70

The study involved pretest (phase 1) and posttest (phase 2) evaluations of students in control and experimental conditions. The total number of correct answers, total number of answers, and difficulty scores are reported for each class and condition.

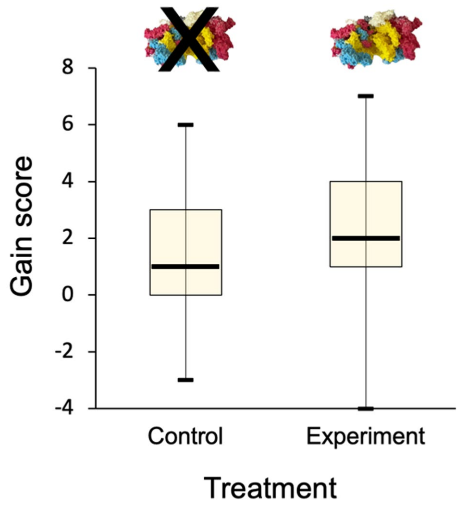

FIGURE 5
Comparison of gain scores obtained from the students who participated in the study (n=total number of students=85). The gain scores were calculated as the differences between the pretest and posttest scores obtained by students within each group. The "Control" group consisted of students that were submitted to the theoretical explanation, and the "Experimental" group consisted of students that were submitted to the presentation of the 3D printed molecules. Horizontal lines indicate median values and boxes indicate 1st and 3rd quartiles. We used the non-parametric Mann–Whitney test for comparison because data distribution was not normal (see text for details).

In conclusion, the method showcased in this article has as a foundation the combination of digital manufacturing processes, which can be used to make (almost) anything, anywhere, with readily-available molecular computational elements (Gershenfeld, 2012). Digital fabrication is a design and production process focused on turning bits into atoms, increasingly present in academic and school settings. On the other hand, databases like PDB are free to use and reunite thousands of digital molecular components, which allows the exportation in formats compatible with 3D printers. Bridging the gap on how simple the process of connecting those two resources is a crucial step to incentivizing teachers to explore it, and also develop systems for their classes in order to help students to better understand the complex concepts underlying the synthetic biology area.

Acknowledgment of any conceptual, methodological, environmental, or material constraints

We had some technical issues during the 3D printing testing phase due to (1) *.stl* files were too large and did not load in the Ultimaker Cura; (2) the settings of the printer were not ideal, and the resulting pieces were low in quality. We solved these issues by reducing the size of the files in Blender and exporting it again as *.stl* files, optimizing the printing supports, and improving the definition of the molecule surfaces.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by CEP/FS-UnB - Comitê de Ética em Pesquisa com Seres Humanos Faculdade de Ciências da Saúde Universidade de Brasília.

The patients/participants provided their written informed consent to participate in this study.

Author contributions

CC conceived the study. HO performed the data collection and adapted the data to 3D printer. MC and AG set up the conditions and printed all the biological structures. CC, MC, AG, AB, SR, LS, and VR elaborated the questionnaire and applied the proposed method in the different classes. HO and EV performed the statistical analysis. HO and CC wrote the manuscript. All authors contributed to the article and approved the submitted version.

Funding

We thank our supporting agencies, DPI/DPG (Decanato de Pesquisa e Inovação e Decanato de Pós-Graduação, from the University of Brasilia, Brazil), DEX (Decanato de Extensão, from the University of Brasilia, Brazil), FAPDF (Research Support Foundation of the Federal District, Brazil), and CNPq (National Council for Scientific and Technological), Brazil.

Acknowledgments

We thank the professors of the Genetic Engineering class, Lídia Maria Pepe de Moraes and Fernando Araripe Gonçalves Torres; the professors of the Molecular Biology classes, Ildinete Silva-Pereira, Marcelo Brígido, and Juliana Franco; and the professor of the Genetics class, Márcio José Poças, for letting us present the 3D printed toggle switch in their classes.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2023.1110464/full#supplementary-material

References

Barone, J., Bayer, C., Copley, R., Barlow, N., Burns, M., Rao, S., et al. (2015). "Nanocrafter: design and evaluation of a DNA nanotechnology game." in *Proceedings of the 10th International Conference on the Foundations of Digital Games (FDG 2015).*

Basics, S. (2009). Serial Cloner. Available at: http://serialbasics.free.fr.Serial_Cloner.html.

Beltramini, L. M., Araújo, A. P., de Oliveira, T. H., Dos Santos Abel, L. D., da Silva, A. R., and Dos Santos, N. F. (2006). A new three-dimensional educational model kit for building DNA and RNA molecules: development and evaluation. *Biochem. Mol. Biol. Educ.* 34, 187–193. doi: 10.1002/bmb.2006.49403403187

Brooks, S. M., and Alper, H. S. (2021). Applications, challenges, and needs for employing synthetic biology beyond the lab. *Nat. Commun.* 12:1390. doi: 10.1038/s41467-021-21740-0

Cameron, D., Bashor, C., and Collins, J. (2014). A brief history of synthetic biology. *Nat. Rev. Microbiol.* 12, 381–390. doi: 10.1038/nrmicro3239

Davenport, J., Pique, M., Getzoff, E., Huntoon, J., Gardner, A., and Olson, A. (2017). A self-assisting protein folding model for teaching structural molecular biology. *Structure* 25, 671–678. doi: 10.1016/j.str.2017.03.001

Diep, P., Boucinha, A., Kell, B., Yeung, B. A., Chen, X., Tsyplenkov, D., et al. (2021). Advancing undergraduate synthetic biology education: insights from a Canadian iGEM student perspective. *Can. J. Microbiol.* 67, 749–770. doi: 10.1139/cjm-2020-0549

El Karoui, M., Hoyos-Flight, M., and Fletcher, L. (2019). Future trends in synthetic biology—a report. *Front. Bioeng. Biotechnol.* 7:175. doi: 10.3389/fbioe.2019.00175

Elowitz, M., and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature* 403, 335–338. doi: 10.1038/35002125

Flores Bueso, Y., and Tangney, M. (2017). Synthetic biology in the driving seat of the bioeconomy. *Trends Biotechnol.* 35, 373–378. doi: 10.1016/j.tibtech.2017.02.002

French, K. E. (2019). Harnessing synthetic biology for sustainable development. *Nat. Sustain.* 2, 250–252. doi: 10.1038/s41893-019-0270-x

Fu, P. (2006). A perspective of synthetic biology: assembling building blocks for novel functions. *Biotechnol. J.* 1, 690–699. doi: 10.1002/biot.200600019

Gardner, T. S., Cantor, C. R., and Collins, J. J. (2000). Construction of a genetic toggle switch in Escherichia coli. *Nature* 403, 339–342. doi: 10.1038/35002131

Gershenfeld, N. (2012). How to make almost anything: the digital fabrication revolution. *Foreign Aff.* 91, 43–57.

Gingold, C., and Douglas, S. M. (2018). Gelbox—an interactive simulation tool for gel electrophoresis. BioRxiv [Preprint]. doi: 10.1101/406132

Gliner, J. A., Morgan, G. A., and Harmon, R. J. (2003). Pretest-posttest comparison group designs: analysis and interpretation. *J. Am. Acad. Child Adolesc. Psychiatry* 42, 500–503. doi: 10.1097/01.CHI.0000046809.95464.BE

Gordy, C. L., Sandefur, C. I., Lacara, T., Harris, F. R., and Ramirez, M. V. (2020). Building the lac operon: a guided-inquiry activity using 3D-printed models. *J. Microbiol. Biol. Educ.* 21:21.1.28. doi: 10.1128/jmbe.v21i1.2091

Goujet, R. (2018). Hero. Coli: A video game empowering stealth learning of synthetic biology: A continuous analytics-driven game design approach. Doctoral dissertation. Sorbonne Paris Cité).

Hansen, A. K., Langdon, T. R., Mendrin, L. W., Peters, K., Ramos, J., and Lent, D. D. (2020). Exploring the potential of 3D-printing in biological education: a review of the literature. *Integr. Comp. Biol.* 60, 896–905. doi: 10.1093/icb/icaa100

Huang, A., Bryan, B., Kraves, S., Alvarez-Saavedra, E., and Stark, J. C. (2022). Implementing hands-on molecular and synthetic biology education using cell-free technology. *Methods Mol. Biol.* 2433, 413–432. doi: 10.1007/978-1-0716-1998-8_25

Huang, H. H., Camsund, D., Lindblad, P., and Heidorn, T. (2010). Design and characterization of molecular tools for a synthetic biology approach towards developing cyanobacterial biotechnology. *Nucleic Acids Res.* 38, 2577–2593. doi: 10.1093/nar/gkq164

Huang, A., Nguyen, P. Q., Stark, J. C., Takahashi, M. K., Donghia, N., Ferrante, T., et al. (2018). BioBits™ explorer: a modular synthetic biology education kit. *Sci. Adv.* 4:eaat5105. doi: 10.1126/sciadv.aat5105

Jacob, F., and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* 3, 318–356. doi: 10.1016/s0022-2836(61)80072-7

Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H., and Phillips, D. C. (1958). A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* 181, 662–666. doi: 10.1038/181662a0

Khalil, A. S., and Collins, J. J. (2010). Synthetic biology: applications come of age. *Nat. Rev. Genet.* 11, 367–379. doi: 10.1038/nrg2775

Kolb, D. A. (1984). *Experiential Learning: Experience as the Source of Learning and Development*. Englewood Cliffs, NJ: Prentice-Hall, 20–38.

Kolb, D. A., Boyatzis, R. E., and Mainemelis, C. (2001). "Experiential learning theory: previous research and new directions" in *Perspectives on Thinking, Learning, and Cognitive Styles*. eds. R. J. Sternberg and L.-F. Zhang (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 227–247.

Koltun, W. L. (1965). Precision space-filling atomic models. *Biopolymers* 3, 665–679. doi: 10.1002/bip.360030606

König, H., Frank, D., Heil, R., and Coenen, C. (2013). Synthetic genomics and synthetic biology applications between hopes and concerns. *Curr. Genomics* 14, 11–24. doi: 10.2174/1389202911314010003

Lineback, J. E., and Jansma, A. L. (2019). PyMOL as an instructional tool to represent and manipulate the myoglobin/hemoglobin protein system. *J. Chem. Educ.* 96, 2540–2544. doi: 10.1021/acs.jchemed.9b00143

Luria, S. E. (1970). The recognition of DNA in bacteria. *Sci. Am.* 222, 88–102. doi: 10.1038/scientificamerican017088

Lv, X., Hueso-Gil, A., Bi, X., Wu, Y., Liu, Y., Liu, L., et al. (2022). New synthetic biology tools for metabolic control. *Curr. Opin. Biotechnol.* 76:102724. doi: 10.1016/j.copbio.2022.102724

Muth, L. T., Jenkins Sánchez, L. R., Claus, S., Salvador Lopez, J. M., and Van Bogaert, I. (2021). A toolbox for digitally enhanced teaching in synthetic biology. *FEMS Microbiol. Lett.* 368:fnab115. doi: 10.1093/femsle/fnab115

Newman, D. L., Stefkovich, M., Clasen, C., Franzen, M. A., and Wright, L. K. (2018). Physical models can provide superior learning opportunities beyond the benefits of active engagements. *Biochem. Mol. Biol. Educ.* 46, 435–444. doi: 10.1002/bmb.21159

Olson, A. J. (2018). Perspectives on structural molecular biology visualization: from past to present. *J. Mol. Biol.* 430, 3997–4012. doi: 10.1016/j.jmb.2018.07.009

Pauling, L., and Corey, R. B. (1951). The pleated sheet, a new layer configuration of polypeptide chains. *Proc. Natl. Acad. Sci. U. S. A.* 37, 251–256. doi: 10.1073/pnas.37.5.251

Pauling, L., Corey, R. B., and Bransson, H. R. (1951). The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci. U. S. A.* 37, 205–211. doi: 10.1073/pnas.37.4.205

Pernaa, J., and Wiedmer, S. (2019). A systematic review of 3D printing in chemistry education – analysis of earlier research and educational use through technological pedagogical content knowledge framework. *Chem. Teach. Int.* 2:20190005. doi: 10.1515/cti-2019-0005

Pinger, C. W., Geiger, M. K., and Spence, D. M. (2019). Applications of 3D-printing for improving chemistry education. *J. Chem. Educ.* 97, 112–117. doi: 10.1021/acs.jchemed.9b00588

Pittalis, M., and Christou, C. (2010). Types of reasoning in 3D geometry thinking and their relation with spatial ability. *Educ. Stud. Math.* 75, 191–212. doi: 10.1007/s10649-010-9251-8

Salzman, M. C., Dede, C., Loftin, R. B., and Chen, J. (1999). A model for understanding how virtual reality aids complex conceptual learning. *Presence Teleop. Virt.* 8, 293–316. doi: 10.1162/105474699566242

Smith, H. O., and Wilcox, K. W. (1970). A restriction enzyme from Hemophilus influenzae. I. Purification and general properties. *J. Mol. Biol.* 51, 379–391. doi: 10.1016/0022-2836(70)90149-x

Stone, B., Kay, D., Reynolds, A., and Brown, D. (2020). 3D printing and service learning: accessible open educational resources for students with visual impairment. *Int. J. Teach. Learn. High. Educ.* 32, 336–346.

UNESCO International Bureau of Education (2019). Exploring STEM competences for the 21st century (current and critical issues in curriculum, learning and assessment, 30). IBE/2019/WP/CD/30 REV. Geneva: IBE.

Wu, H.-K., Krajcik, J. S., and Soloway, E. (2001). Promoting understanding of chemical representations: Students' use of a visualization tool in the classroom. *J. Res. Sci. Teach.* 38, 821–842. doi: 10.1002/tea.1033

Zientek, L., Nimon, K., and Hammack-Brown, B. (2016). Analyzing data from a pretest-posttest control group design: the importance of statistical assumptions. *Eur. J. Train. Dev.* 40, 638–659. doi: 10.1108/EJTD-08-2015-0066

TYPE Review
PUBLISHED 31 March 2023
DOI 10.3389/feduc.2023.1163024

OPEN ACCESS

EDITED BY
Wang-Kin Chiu,
The Hong Kong Polytechnic University,
China

REVIEWED BY
Sheila L. Macrine,
University of Massachusetts Dartmouth,
United States
Rafidah Abd Karim,
MARA University of Technology,
Malaysia

*CORRESPONDENCE
Ard W. Lazonder
✉ ard.lazonder@ru.nl

†PRESENT ADDRESS
Sifra E. Muilwijk,
Institute of Movement Studies,
University of Applied Sciences Utrecht,
Utrecht,
Netherlands

SPECIALTY SECTION
This article was submitted to
STEM Education,
a section of the journal
Frontiers in Education

RECEIVED 10 February 2023
ACCEPTED 03 March 2023
PUBLISHED 31 March 2023

CITATION
Muilwijk SE and Lazonder AW (2023) Learning from physical and virtual investigation: A meta-analysis of conceptual knowledge acquisition. *Front. Educ.* 8:1163024.
doi: 10.3389/feduc.2023.1163024

Learning from physical and virtual investigation: A meta-analysis of conceptual knowledge acquisition

Sifra E. Muilwijk[1†] and Ard W. Lazonder[2*]

[1]School of Pedagogical and Educational Sciences, Radboud University, Nijmegen, Netherlands,
[2]Behavioural Science Institute, Radboud University, Nijmegen, Netherlands

Should students investigate with tangible objects and apparatus or are digitally simulated materials and equipment an adequate or perhaps even preferred alternative? This question remains unanswered because empirical evidence is inconclusive and previous reviews are descriptive and synthesize a limited number of studies with small samples. This meta-analysis, therefore, assessed the relative effectiveness of physical versus virtual investigation in terms of conceptual knowledge acquisition and examined whether and how the aggregate effect size was moderated by substantive and methodological study features. Following a systematic search of Web of Science and ERIC for the period 2000–2021, 35 studies comparing physical and virtual investigations were selected for inclusion. Hedges' g effect sizes for conceptual knowledge acquisition were computed and analyzed using a random effects model. The results showed no overall advantage of either mode of investigation (g=–0.14, 95% CI [–0.33, 0.06]). However, moderator analysis indicated that virtual investigation is more effective for adults compared with adolescents and children, and when touching objects or equipment does not provide relevant sensory information about the concept under study. These results imply that STEM teachers can decide for themselves whether to opt for physical or virtual investigation except when teaching adult students or when touch sensory feedback is substantively irrelevant; in those cases, virtual investigation is preferable.

KEYWORDS
virtual labs, simulation – computers, physicality, manipulatives and experimentation, touch

1. Introduction

Technological advancements have significantly extended the opportunities to include investigations in courses for students of all ages. The past decade has witnessed several successful initiatives to grant teachers of science, technology, engineering, and mathematics (STEM) free access to computer simulations and online laboratories for teaching and learning in K12 classrooms and beyond. Although designers of these technologies are optimistic about the value of virtual investigation for learning (Perkins et al., 2012; De Jong et al., 2014), empirical evidence is typically mixed. Some studies confirm that virtual investigation is more effective than physical investigation (e.g., Chao et al., 2016), whereas other studies found the opposite effect (e.g., Zacharia et al., 2012) or report no differences (e.g., Renken and Nunez, 2013). As this body of research has, to the best of our knowledge, not been quantitatively reviewed, the true virtue of virtual investigation is yet unknown. This observation sparked the idea for this meta-analysis, which aimed to examine the relative effectiveness of physical versus virtual investigations.

In line with Klahr et al. (2007), we use the term physical investigation to refer to hands-on inquiries where students interact with tangible objects and equipment to acquire a conceptual understanding of the topic being studied. Although such investigations admittedly enable students to strengthen their research skills, learn to collaborate with peers, and build interest in STEM-related careers, these learning outcomes were outside the scope of this meta-analysis. Virtual investigation, then, is the digital analog of physical investigation in that students' examinations involve simulated material and apparatus provided by a computer simulation, virtual laboratory, or virtual–reality application. Both definitions are fleshed out more in the sections below. Following a short overview of instructional approaches that incorporate student investigations, we zoom in on the unique affordances of physical and virtual investigation and summarize the results of previous narrative reviews that contrasted these modes of investigation.

2. Theoretical foundation

2.1. Learning through investigation

Preschoolers learn from play and by exploring the world around them. Schools respond to this investigative drive by engaging children in inquiry projects, for instance, to examine how long it takes for colored ink to dissolve in hot and cold water. High school students spend quite some time in the school science laboratory, and throughout higher education, student research progressively approximates authentic scientific practices. All these instances are rooted in the long-standing belief that the act of investigating is productive to learning because finding things out by oneself leads to more meaningful and sustainable knowledge than being told by a teacher (Dewey, 1900; Schneider et al., 2022; De Jong et al., in press).

Student investigations are integral to instructional approaches such as experiential learning, problem-based learning, and inquiry-based learning. Although generally embraced by policymakers and field experts, some educational scientists have challenged the effectiveness of these approaches based on a lack of teacher guidance (e.g., Kirschner et al., 2006; Zhang et al., 2022). A comprehensive meta-analysis confirmed that learning through unguided investigation is less effective than explicit instruction. However, students who were guided during their investigation learned more than students who studied the same material through expository methods (Alfieri et al., 2011). In other words, instructional approaches that include student investigations are effective if adequate guidance is provided.

Which type of guidance is appropriate for which types of learners is still debated. Contrary to the intuitive belief that young learners need more specific guidance than older learners do, student age does not moderate the influence of guidance on learning activities and learning outcomes (Lazonder and Harmsen, 2016). A plausible explanation might be that the complexity of students' investigations increases with age: older students not only examine more difficult topics but are also exposed to more open forms of inquiry that necessitate specific forms of guidance (Bell et al., 2005). As such, creating effective learning arrangements that include student research is a balancing act that becomes even more challenging if teachers can choose between physical and virtual investigations.

2.2. The case for the physical investigation

Imagine giving a child a set of cubes and spheres to investigate what determines how fast objects sink in water. While experimenting, the child receives sensory feedback through the eyes (e.g., spheres sink in a straight line whereas cubes whirl down) and ears (sound indicates when an object hits the bottom of the water cylinder). A unique additional affordance of physical investigations is that handling the spheres and cubes produces touch sensory feedback about their mass and surface not normally available in virtual investigations. [It can be mimicked by a haptic device, but these tools are still rare in educational settings (Luo et al., 2021) and were, therefore, not included in this meta-analysis].

The educational importance of touch sensory feedback is articulated in theories of embodied cognition, which assert that a person interacting with the material world creates 'embodied' knowledge of physical objects and phenomena (e.g., Gallese and Lakoff, 2005; Barsalou, 2008). Neuroimaging studies have provided evidence that conceptual understanding is stored in the sensory–motor circuits of the brain, meaning that the brain regions involved in seeing, hearing, and touching objects are also activated during recall (Kiefer and Trumpp, 2012). If students are deprived of touch, conceptual knowledge becomes less rich as it is exclusively based on verbal and auditory stimuli (Zacharia, 2015). The child in our sinking objects example learns about the mass of the objects by picking them up and the experience of feeling the difference between, for instance, a 10 and 100 g cube complements mass-related information from the other senses.

While embodied cognition theories emphasize the storage and retrieval of information, the additional sensory channel theory addresses the encoding process. Rooted in theoretical conceptions of working memory (Baddeley, 2012) and cognitive load (Sweller et al., 2019), this theory postulates that the brain has separate processing channels for visual, auditory, and tactile information. If multiple channels are used for learning a particular piece of information, the effective working memory capacity expands and, hence, the chance of better learning outcomes increases. Note that this theory merely applies to sensory information relevant to the concept students are investigating. Suppose the child in our example senses that metal objects feel colder than Teflon objects, then this tactile feedback will not lower her cognitive load when examining how the shape of an object influences its sink time.

2.3. The case for the virtual investigation

Proponents of virtual investigation point to the practical advantages of simulations and virtual laboratories. These digital environments require little preparation from the teacher and enable students to design and conduct many investigations in a short amount of time (De Jong et al., 2013). Virtual investigations also offer a viable alternative for physical investigation if material or apparatus is expensive or when the research site is geographically remote (Hannel and Cuevas, 2018)—think, for example, of a field trip to the Falkland telescope to have astronomy students observe distant galaxies. Similar advantages apply when the topic of investigation is rare (e.g., lunar eclipses) or dangerous (e.g., radioactivity).

The learning benefits of virtual investigation are essentially twofold. Contentwise, designers of digital investigation environments can impose productive constraints on students—for instance, by simplifying a phenomenon or restricting the values that can be set in an experiment—or provide visualizations that allow students to perceive what is not directly observable in the material world (De Jong et al., 2013; Sullivan et al., 2017). These options aim to reduce intrinsic cognitive load. Extraneous cognitive load can be decreased by embedding instructional support features in the virtual environment, which is the second learning advantage. Software designers can, for instance, use virtual–reality technology to direct students' attention to important parts of the screen at key moments during an inquiry or augment digital objects and processes with additional explanations (De Jong et al., 2013).

2.4. Previous narrative reviews

An early research overview by Ma and Nickerson (2006) found no significant and consistent difference between physical and virtual investigation. De Jong et al. (2013) reached a similar conclusion and speculated about possible differential effects by suggesting that young learners might benefit more from physical investigation because they tend to lack tactile experience with the objects or processes under study. Virtual investigation, according to De Jong et al., might be more advantageous in situations that align with the learning advantages described in the previous paragraph. Zacharia (2015) also concluded that touch sensory feedback from physical investigation is not a requirement for the acquisition of conceptual knowledge.

Related literature overviews challenge the latter conclusion by showing that haptic augmentation of virtual environments often improves the development of conceptual knowledge (Minogue and Jones, 2006; Zacharia, 2015) and procedural skills (Rangarajan et al., 2020). This positive trend seems due to the fact that all haptic devices provided 'force feedback' directly relevant to the topic being studied (e.g., gears and lever principles). Future research could test this presumption by comparing physical and virtual investigation, the latter without haptics, in situations where touch sensory information helps students build an understanding of the concepts or processes they are investigating. Furthermore, a meta-analysis in the field of mathematics education strengthened the tentative conclusion regarding the moderating influence of learners' age (Carbonneau et al., 2013). Using concrete manipulatives in math classes was found to be more effective for children in the concrete operational stage compared with children in the formal operational stage, allegedly because younger children rely more on physical interaction with the material world when constructing meaning than older children who are capable of formal operational reasoning.

In summary, previous research integrations converge on the equivalent effectiveness of physical and virtual investigation but differ regarding the existence of possible age-related differences as well as the educational affordances of touch. However, as most of these works descriptively synthesized a selective number of studies with small sample sizes, more rigorous and quantitative research integrations are needed to draw any definitive conclusion.

3. Research questions

This meta-analysis aimed to answer three research questions:

1. What is the relative effectiveness of physical versus virtual investigation in terms of conceptual knowledge acquisition?
2. How does this relative effectiveness depend on the substantive contribution of touch sensory information to the concept under study?
3. How does this relative effectiveness depend on the students' age?

4. Method

4.1. Search and selection of studies

The literature was searched for studies that satisfied the following inclusions criteria:

1. The study examined students investigating STEM-related topics for learning purposes.
2. The study compared the conceptual knowledge acquisition of students who did their research with physical materials to that of students who performed the same investigation with virtual materials.
3. The study was set up to ensure that similar instructional regimes were implemented in both conditions.
4. The study controlled for possible differences in prior domain knowledge either by randomization or analyzing pre- and post-assessment scores.
5. The study administered a between-subject design and reported data from which effect sizes can be calculated.
6. The study was published between 2000 and 2021 and is available online in full text.

The search and selection processes are visualized in Figure 1. All searches were performed in the Web of Science Core Collection and the ERIC repository using the following query: [(physical experiment* OR physical lab* OR hands-on) AND (virtual experiment* OR virtual lab* OR hands-off OR simulation*)]. The Web of Science search, restricted to the SSCI and SCI Expanded citation indexes and further limited to the categories 'Education Educational Research,' 'Education Scientific Disciplines,' 'Psychology Experimental,' and 'Psychology Multidisciplinary,' uncovered 739 reports. The ERIC database was searched similarly except that all search terms were wrapped in quotes to perform a literal search, and the results were limited to publications available in full text. This search returned 87 hits. Next, all 826 reports were retrieved and subjected to a title and abstract screening. A total of 58 reports passed this initial test, and after one duplicate was removed, 57 reports were read in full to assess their eligibility for inclusion. In addition, the perusal of the citations in previous reviews (De Jong et al., 2013; Zacharia, 2015) yielded eight reports that were not identified through the online search. These reports were retrieved and screened similarly, which led to the inclusion of one additional report. This brought the total number of included reports to 34. As one of these investigations presented data from two experiments with separate samples, the total number of studies in this meta-analysis was 35. Their main characteristics are summarized in Supplementary Appendix 1.

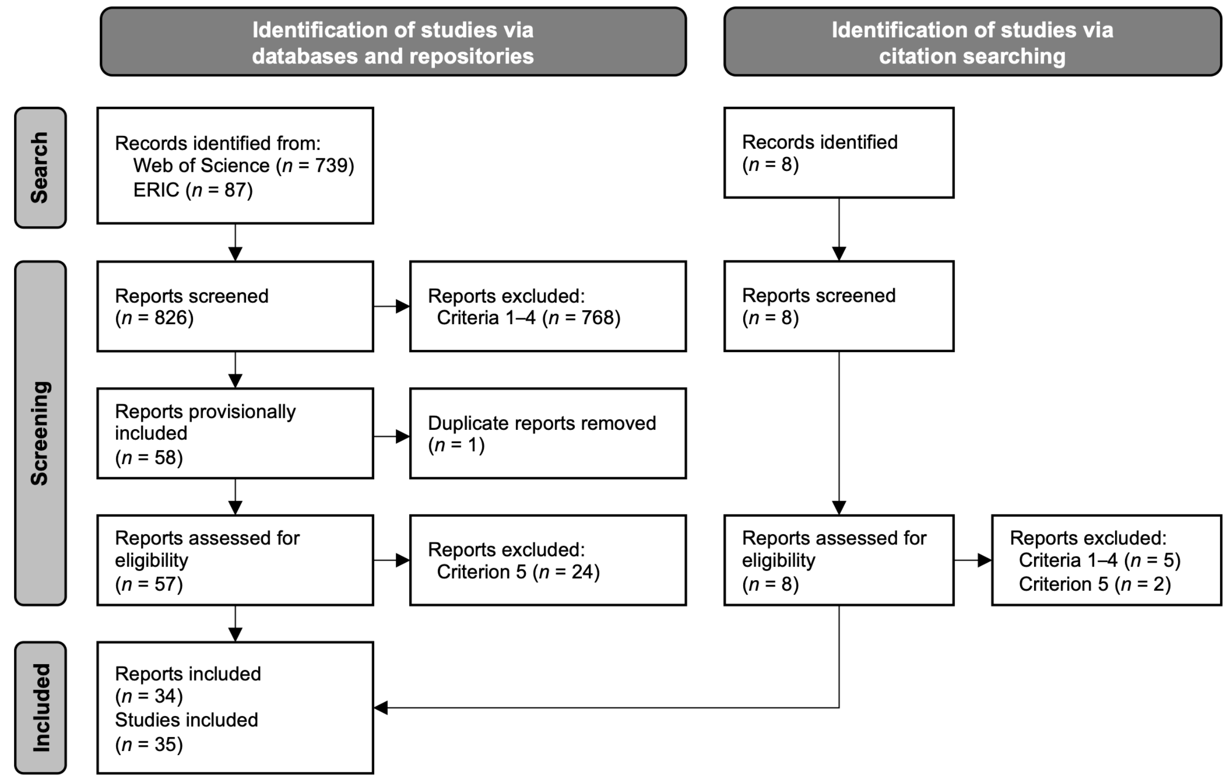

FIGURE 1
PRISMA flow diagram of the study search and selection process.

4.2. Coding of moderator and outcome variables

Students' *knowledge acquisition* served as the main outcome measure. It was defined as the conceptual knowledge participants developed through either physical or virtual investigation, as indicated by assessments administered during or shortly after the learning process. The first moderator, *assessment type*, classified the measurement used as a multiple-choice test, constructed-response test, performance-based assessment, or a combined format.

The next two moderators served to answer this study's research questions. The first one, *tactile feedback*, indicated whether physical manipulation provided touch sensory information that helps students build an understanding of the concepts they are investigating. The moderator *student age* gave a broad indication of the sample's mean age. In keeping with Piaget's and Erikson's stages of cognitive development (Thomas, 2005), a distinction was made between school-age children (6–11 years), adolescents (12–18 years), and young adults (19–27 years). In case participants' age was not provided, the age category was inferred from the students' grade levels, considering the differences in educational systems across countries.

The remaining moderators provided some descriptive details of the included studies. *Publication year* was used as an approximation of the time when the study was conducted. As computer technology becomes increasingly more sophisticated, older studies using computer simulations might yield different results than recent research with highly advanced virtual investigation facilities. To investigate whether such a differential effect exists, studies were classified according to the decade of publication (2000–2010 or 2011–2021). *The research setting* concerned the site where the study took place. Two broad categories were distinguished: research laboratory and regular classroom. The former indicated that data were collected in a researcher-controlled environment, for example, a genuine university research laboratory or a separate room in the school building. Studies that were carried out in an authentic learning environment (e.g., a lecture room, the school's science laboratory, or a computer laboratory) were placed in the category 'regular classroom.'

Interrater agreement was determined in case moderator coding required subjective interpretation by the raters. Agreement on 'assessment type' (78%, Fleiss' $\kappa = 0.71$) and 'tactile feedback' (86%, Fleiss' $\kappa = 0.72$) was substantial according to the benchmarks proposed by Landis and Koch (1977). All disagreements were resolved through discussion. The remaining moderators were coded by the first author, who conferred with the second author when in doubt. Fisher's exact tests were run to determine whether the five moderators were related. Using a Bonferroni-corrected alpha level of 0.005, none of the comparisons turned out to be statistically significant, the p-values were >0.133, which means that all moderators were mutually independent.

4.3. Computation of effect sizes

Standardized mean differences were computed and corrected for upward small-sample bias. This effect size metric, known as Hedges' g, was calculated as follows:

$$g = \frac{M_1 - M_2}{SD_{pooled}} \times \left(\frac{N-3}{N-2.25} \right) \times \sqrt{\frac{N-2}{N}}$$

where N is the total sample size, M_1 is the mean knowledge gain score of the students in the physical investigation condition, M_2 is the mean gain of the students in the virtual investigation condition, and SD is the weighted standard deviation of both groups combined. If gain scores were not reported, the study's effect size was calculated from pre- and post-assessment scores, test statistics (F, t, and χ^2), or frequency distributions, using the conversion formulas by Lipsey and Wilson (2001). Note that, as per computation, a positive effect size indicates that students learned more from physical investigation, whereas a negative effect size denotes higher learning from virtual investigation.

Studies reporting data for multiple subgroups or multiple outcome measures were handled according to the guidelines proposed by Borenstein et al. (2009). Specifically, one study presented separate scores for the high and low achievers in both the physical and virtual investigation conditions. To reduce bias, scores of the two cohorts within each condition were combined to yield a summary effect. Other studies assessed students' knowledge acquisition by multiple post-tests. As these tests were equally relevant in determining which concepts students had learned through investigation, their scores were combined to compute the study's effect size.

4.4. Data analysis

Main analyses were conducted with Meta Essentials (Suurmond et al., 2017). The random effects model was used because studies examining different-aged students engaged in different inquiry tasks with different objects and equipment are unlikely to share the same true effect size. Following a descriptive analysis of the studies' effect sizes, the summary effect was tested for significance by a z-test. Egger's regression test (Egger et al., 1997) and Orwin's (1983) Fail-safe N were used to determine whether and to what extent the observed overall effect was subject to publication bias. Next, Q-tests based on analysis of variance were used to determine whether the between-study variation in effect sizes was attributable to the moderator variables. If so and where appropriate, planned comparisons (Hedges and Pigott, 2004) were made to unveil which moderator categories differed significantly from one another.

5. Results

Data for this meta-analysis were extracted from 35 studies with 3,303 participants. The effect sizes of two studies were significantly greater than zero ($g = 1.23$ and 1.63), which denotes a benefit of physical investigation over virtual investigation. Seven studies had a significant negative effect size in the range of −1.51 to −0.45, which indicates in favor of virtual investigation, and in 26 studies, the physical–virtual comparison was a tie ($-0.37 < g < 0.52$).

The studies' overall mean effect size (g) was −0.14, $SE = 0.09$, 95% CI [−0.33, 0.06]. The I^2 statistic indicated that 82.44% of the effect size variance reflects true score variation; the variance of the true effect size (τ^2) was 0.21. As can be inferred from the confidence interval, the investigation mode had no significant overall effect on students' knowledge acquisition, $z = -1.43$, $p = 0.152$, meaning that experimenting with physical and virtual materials is equally beneficial to concept learning. Egger's regression test showed no sign of publication bias as the estimated intercept (−1.81) did not differ significantly from zero, $t(34) = 0.73$, $p = 0.472$. Orwin's Fail-safe N indicated that 476 studies with a nil effect would be needed to turn the Hedges' g to zero.

The results of the moderator analyses showed that the variation in effect sizes was independent of how students' conceptual knowledge was assessed and in which year a study was published (see Table 1). However, a significant moderating effect was found for tactile feedback. This result indicates that physical and virtual investigations yield comparable knowledge gains if touching materials provide relevant information about the concepts to be learned, but that virtual investigation is more effective when the touch experience is extraneous.

The participants' age also moderated the findings. The mean effect size of studies conducted with children was higher than that of studies with adolescents and adults combined, $z = 3.58$, $p < 0.001$, and the difference between the latter two age groups was also significant, $z = 2.36$, $p = 0.009$. The confidence intervals in Table 1 further show that adults benefit more from virtual investigations than physical investigations, whereas children and adolescents benefit as much from either mode of investigation.

The summary effect also depended on the site where a study was carried out. Studies conducted under researcher-controlled circumstances had a significantly higher mean effect size than studies performed in more authentic settings such as a regular classroom. The direction of these effect sizes implies that students benefit more from virtual investigation if their research is situated in authentic settings guided by regular classroom teachers. But when students' inquiry takes place in a quiet space under the surveillance of a proctor, physical investigation is more effective than virtual investigation. It should be noted that the distribution of studies among these two categories was rather skewed and may have impacted the findings.

6. Discussion

The summary effect of the 35 primary studies included in this meta-analysis indicates that physical and virtual investigation are generally equally effective in promoting students' conceptual knowledge of STEM-related topics. This outcome confirms the tentative conclusion from descriptive reviews (Ma and Nickerson, 2006; De Jong et al., 2013; Zacharia, 2015) and implies that the true effect, although slightly in favor of virtual investigation, is close to zero. The fact that this result was independent of the year in which a study was published further suggests that technological advancements have no impact on how much knowledge students acquire from virtual investigation relative to physical investigation. In other words, computer simulations from the early 2000s are as productive to concept learning as contemporary virtual laboratories with highly realistic 3D rendering.

However, the equivalence of investigation modes does not apply to all learning situations. Adults, for example, benefit more from virtual investigation than physical investigation, while no such

TABLE 1 Results of the moderator analyses.

	k	*N*	*g*	95% CI	*Q*	*p*	I^2
Assessment type					2.48	0.479	0.00
Multiple choice	12	996	−0.20	[−0.42, 0.02]			
Constructed response	16	1,119	−0.02	[−0.37, 0.32]			
Combined	3	336	−0.04	[−0.26, 0.17]			
Performance based	4	852	−0.42	[−0.99, 0.15]			
Tactile feedback					4.51	0.034	77.83
Relevant	15	1,334	0.09	[−0.24, 0.42]			
Irrelevant	20	1969	−0.30	[−0.49, −0.11]			
Student age					8.79	0.012	88.62
Children	10	801	0.22	[−0.25, 0.69]			
Adolescents	8	639	−0.04	[−0.17, 0.08]			
Adults	17	1863	−0.38	[−0.59, −0.17]			
Publication year					0.01	0.967	0.00
2000–2010	8	556	−0.13	[−0.36, 0.10]			
2011–2021	27	2,747	−0.14	[−0.37, 0.10]			
Research setting					6.18	0.013	83.82
Research lab	4	220	0.81	[0.01, 1.61]			
Classroom	31	3,083	−0.24	[−0.39, −0.09]			

benefit was found in adolescents and children. Whether the comparable effectiveness of physical and virtual investigation for younger learners is attributable to their developmental stage or a lack of experience with the objects being investigated (De Jong et al., 2013) cannot be concluded from our meta-analysis. Theoretical evidence supports the former option, but the observed superiority of virtual investigation in adults supports the latter. Future research could resolve this discord by comparing the knowledge gains of children and adults in an investigation with familiar and unfamiliar tangible objects.

The relative effectiveness of investigation modes also depends on the substantive contribution of tactile feedback. Virtual experimentation is more effective when the focal variables in an investigation cannot be experienced by touch. But when tactile cues do provide relevant information, physical investigation is just as effective. This differential effect is in line with the additional channel theory, which assumes that information from touch is processed in a distinct part of the human brain and, hence, reduces cognitive load during learning. A direct assessment of the latter claim could unfortunately not be made here because none of the included studies measured students' cognitive load—which is remarkable because quite many studies mentioned cognitive load reduction as one of the advantages of either physical or virtual investigation.

Implications for theories of embedded cognition are less straightforward. On the one hand, our results lend no direct support to the notion that physical manipulation promotes conceptual understanding since manipulating virtual objects on a computer screen was generally equally effective. On the other hand, the data do not disqualify the embodiment idea either because the physical experience of touch could have compensated for the absence of the affordances of virtual investigation, such as simplifying, annotating, and visualizing concepts and processes. The value of such features was demonstrated by Lee et al. (2006), who found that separated screen displays and optimized visual representations enhance middle school students' conceptual understanding.

Virtual experimentation environments can also be augmented by haptic feedback. Previous reviews have shown that incorporating haptics can produce significant gains in students' conceptual knowledge (Minogue and Jones, 2006; Zacharia, 2015), but its implementation in educational research and practice is still in its infancy (Luo et al., 2021). Of the few studies we found, none satisfied the inclusion criteria so the comparison between physical investigation with haptic-augmented virtual investigation is yet to be made. Once the body of research has grown, it would be interesting to replicate this meta-analysis and focus specifically on the facilitative role of haptic feedback, in particular when touch conveys information relevant to conceptual understanding.

Beyond contrasting physical and virtual investigations, scholars have started to consider how the two are best combined. The conclusions are still indecisive as some studies favored the physical–virtual sequence (e.g., Winn et al., 2006), other studies the reverse order (e.g., Toth et al., 2014) or reported no difference between both sequences (Flegr et al., 2023). The results of our meta-analysis suggest that starting with physical investigation is preferred when students have an insufficient tactile experience with the concepts or materials being studied, which is often the case with children. When virtual investigation precedes physical investigation, students can benefit from the unique affordances of virtual investigation to efficiently acquire basic knowledge and then deepen and broaden this understanding by investigating the same concepts in more authentic (i.e., 'messy') physical contexts. Although our results provide no direct implications for this option, it seems appropriate for use with adolescents and adults.

Several limitations should be considered when interpreting the results of our meta-analysis. One constraining factor is that the number of included studies was quite small and disproportional across STEM domains. In total, 23 of the 35 studies (66%) were carried out in physics classes, so the current findings do not necessarily apply to other domains. In a similar vein, very few included studies assessed learning outcomes beyond conceptual knowledge, such as students' inquiry skills or their understanding of the nature of science. With a larger set of studies, these outcome measures could have been analyzed to paint a more complete picture of the relative effectiveness of physical and virtual investigations. On a related matter, non-cognitive learning outcomes such as student motivation could be examined to establish whether students are equally interested in doing physical and virtual investigations. Finally, our meta-analysis did not attend to the role of the teacher. This leaves questions regarding whether teachers guide their students equally well during physical and virtual investigations. Research answering questions like these could provide valuable explanations for the relative effectiveness of both modes of investigation.

To conclude, although physical and virtual investigation are generally equally beneficial to promote students' conceptual understanding, the virtual variant is preferred when students are over 18 and have to investigate concepts for which tactile feedback is substantively irrelevant. We, therefore recommend university teachers and adult educators to let students investigate with virtual material and equipment, in particular when the research is conducted in regular classrooms; a switch to physical investigations can be considered if tactile feedback provides relevant sensory information about the concepts being studied or when students conduct their investigation under well-controlled circumstances. Elementary school and high school teachers can decide on a case-by-case basis whether to opt for physical or virtual investigation. They can base their choice on personal preferences and pragmatic considerations while bearing in mind that virtual investigation is more effective when it is not possible to work one-on-one with individual or small groups of students.

Data availability statement

The dataset analyzed for this study is available upon request to the corresponding author.

Author contributions

SM: conceptualization, methodology, systematic search and coding of studies, and writing—reviewing and editing. AL: conceptualization, methodology, validation, formal analysis, writing—original draft, visualization, and supervision. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2023.1163024/full#supplementary-material

References

References marked with an asterisk (*) indicate studies included in the meta-analysis.

* Abdel-Maksoud, N. F. (2018). When virtual becomes better than real: Investigating the impact of a networking simulation on learning and motivation. *Int. J. Educ. Pract.* 6, 253–270. doi: 10.18488/journal.61.2018.64.253.270

* Ajredini, F., Izairi, N., and Zajkov, O. (2013). Real experiments versus PhET simulations for better high-school students' understanding of electrostatic charging. *Eur J Phys Educ* 5, 59–70. doi: 10.20308/ejpe.v5i1.63

Alfieri, L., Brooks, P. J., Aldrich, N. J., and Tenenbaum, H. R. (2011). Does discovery based instruction enhance learning? *J. Educ. Psychol.* 103, 1–18. doi: 10.1037/a0021017

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annu. Rev. Psychol.* 63, 1–29. doi: 10.1146/annurev-psych-120710-100422

* Baki, A., Kosa, T., and Guven, B. (2011). A comparative study of the effects of using dynamic geometry software and physical manipulatives on the spatial visualisation skills of pre-service mathematics teachers. *Br. J. Educ. Technol.* 42, 291–310. doi: 10.1111/j.1467-8535.2009.01012.x

* Barrett, R., Gandhi, H. A., Naganathan, A., Daniels, D., Zhang, Y., Onwunaka, C., et al. (2018). Social and tactile mixed reality increases student engagement in undergraduate lab activities. *J. Chem. Educ.* 95, 1755–1762. doi: 10.1021/acs.jchemed.8b00212

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645. doi: 10.1146/annurev.psych.59.103006.093639

Bell, R. L., Smetana, L., and Binns, I. (2005). Simplifying inquiry instruction. *Sci. Teach.* 72, 30–33.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., and Rothstein, H. R. (2009). *Introduction to meta-analysis*. Hoboken, NJ: Wiley.

* Brown, S. E. (2007). *Counting blocks or keyboards? A comparative analysis of concrete versus virtual manipulatives in elementary school mathematics concepts*. (Unpublished Master's Thesis). Detroit (MI): Marygrove College.

Carbonneau, K. J., Marley, S. C., and Selig, J. P. (2013). A meta-analysis of the efficacy of teaching mathematics with concrete manipulatives. *J. Educ. Psychol.* 105, 380–400. doi: 10.1037/a0031084

* Chang, W. L., Yuan, Y., Lee, C. Y., Chen, M. H., and Huang, W. G. (2013). Using magic board as a teaching aid in third grader learning of area concepts. *Educ. Technol. Soc.* 16, 163–173.

Chao, J., Chiu, J. L., DeJaegher, C. J., and Pan, E. A. (2016). Sensor-augmented virtual labs: Using physical interactions with science simulations to promote understanding of gas behavior. *J. Sci. Educ. Technol.* 25, 16–33. doi: 10.1007/s10956-015-9574-4

* Chen, S., Chang, W. H., Lai, C. H., and Tsai, C. Y. (2014). A comparison of students' approaches to inquiry, conceptual learning, and attitudes in simulation-based and microcomputer-based laboratories. *Sci. Educ.* 98, 905–935. doi: 10.1002/sce.21126

* Chien, K. P., Tsai, C. Y., Chen, H. L., Chang, W. H., and Chen, S. (2015). Learning differences and eye fixation patterns in virtual and physical science laboratories. *Comput. Educ.* 82, 191–201. doi: 10.1016/j.compedu.2014.11.023

* Chini, J. J., Madsen, A., Gire, E., Rebello, N. S., and Puntambekar, S. (2012). Exploration of factors that affect the comparative effectiveness of physical and virtual

manipulatives in an undergraduate laboratory. *Phys. Rev. Spec. Top. Phys. Educ. Res.* 8:010113. doi: 10.1103/PhysRevSTPER.8.010113

* Darrah, M., Humbert, R., Finstein, J., Simon, M., and Hopkins, J. (2014). Are virtual labs as effective as hands-on labs for undergraduate physics? A comparative study at two major universities. *J. Sci. Educ. Technol.* 23, 803–814. doi: 10.1007/s10956-014-9513-9

De Jong, T., Lazonder, A. W., Chinn, C. A., Fischer, F., Gobert, J., Hmelo-Silver, C., et al. (in press). *Let's talk evidence – The case for combining inquiry-based and direct instruction*. Educational Research Review.

De Jong, T., Linn, M. C., and Zacharia, Z. C. (2013). Physical and virtual laboratories in science and engineering education. *Science* 340, 305–308. doi: 10.1126/science.1230579

De Jong, T., Sotiriou, S., and Gillet, D. (2014). Innovations in STEM education: The go-lab federation of online labs. *Smart Learn Environ.* 1:3. doi: 10.1186/s40561-014-0003-6

Dewey, J. (1900). *The school and society*. Chicago, IL: The University of Chicago Press.

Egger, M., Smith, G. D., Schneider, M., and Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *Br. Med. J.* 315, 629–634. doi: 10.1136/bmj.315.7109.629

* Ekmekci, A., and Gulacar, O. (2015). A case study for comparing the effectiveness of a computer simulation and a hands-on activity on learning electric circuits. *EURASIA J. Math. Sci. Technol. Educ.* 11, 765–775. doi: 10.12973/eurasia.2015.1438a

Flegr, S., Kuhn, J., and Scheiter, K. (2023). When the whole is greater than the sum of its parts: Combining real and virtual experiments in science education. *Comput. Educ.* 197:104745. doi: 10.1016/j.compedu.2023.104745

Gallese, V., and Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cogn. Neuropsychol.* 22, 455–479. doi: 10.1080/02643290442000310

* Gecu-Parmaksiz, Z., and Delialioglu, Ö. (2019). Augmented reality-based virtual manipulatives versus physical manipulatives for teaching geometric shapes to preschool children. *Br. J. Educ. Technol.* 50, 3376–3390. doi: 10.1111/bjet.12740

* Gibbard, L. L., and Salajan, F. (2009). A novel interactive online module in a traditional curriculum through a blended learning approach. *Electr. J. E-Learning* 7, 301–308.

* Hannel, S. L., and Cuevas, J. (2018). A study on science achievement and motivation using computer-based simulations compared to traditional hands-on manipulation. *Georgia Educ. Res.* 15, 40–55. doi: 10.20429/ger.2018.15103

* Hawkins, I., and Phelps, A. J. (2013). Virtual laboratory vs. traditional laboratory: Which is more effective for teaching electrochemistry? *Chem. Educ. Res. Pract.* 14, 516–523. doi: 10.1039/C3RP00070B

Hedges, L. V., and Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychol. Methods* 9, 426–445. doi: 10.1037/1082-989X.9.4.426

* Hensen, C., Glinowiecka-Cox, G., and Barbera, J. (2020). Assessing differences between three virtual general chemistry experiments and similar hands-on experiments. *J. Chem. Educ.* 97, 616–625. doi: 10.1021/acs.jchemed.9b00748

* Husnaini, S. J., and Chen, S. (2019). Effects of guided inquiry virtual and physical laboratories on conceptual understanding, inquiry performance, scientific inquiry self-efficacy, and enjoyment. *Phys. Rev. Phys. Educ. Res.* 15:010119. doi: 10.1103/PhysRevPhysEducRes.15.010119

* Jaakkola, T., and Nurmi, S. (2008). Fostering elementary school students' understanding of simple electricity by combining simulation and laboratory activities. *J. Comput. Assist. Learn.* 24, 271–283. doi: 10.1111/j.1365-2729.2007.00259.x

* Kapici, H. O., Akcay, H., and De Jong, T. (2019). Using hands-on and virtual laboratories alone or together—which works better for acquiring knowledge and skills? *J. Sci. Educ. Technol.* 28, 231–250. doi: 10.1007/s10956-018-9762-0

Kiefer, M., and Trumpp, N. M. (2012). Embodiment theory and education: The foundations of cognition in perception and action. *Trends Neurosci. Educ.* 1, 15–20. doi: 10.1016/j.tine.2012.07.002

Kirschner, P. A., Sweller, J., and Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educ. Psychol.* 41, 75–86. doi: 10.1207/s15326985ep4102_1

Klahr, D., Triona, L. M., and Williams, C. (2007). Hands on what? The relative effectiveness of physical versus virtual materials in an engineering design project by middle school children. *J. Res. Sci. Teach.* 44, 183–203. doi: 10.1002/tea.20152

Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174. doi: 10.2307/2529310

* Lazonder, A. W., and Ehrenhard, S. (2014). Relative effectiveness of physical and virtual manipulatives for conceptual change in science: How falling objects fall. *J. Comput. Assist. Learn.* 30, 110–120. doi: 10.1111/jcal.12024

Lazonder, A. W., and Harmsen, R. (2016). Meta-analysis of inquiry-based learning: Effects of guidance. *Rev. Educ. Res.* 86, 681–718. doi: 10.3102/0034654315627366

Lee, H., Plass, J. L., and Homer, B. D. (2006). Optimizing cognitive load for learning from computer-based science simulations. *J. Educ. Psychol.* 98, 902–913. doi: 10.1037/0022-0663.98.4.902

Lipsey, M. W., and Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Luo, H., Li, G., Feng, Q., Yang, Y., and Zuo, M. (2021). Virtual reality in K-12 and higher education: A systematic review of the literature from 2000 to 2019. *J. Comput. Assist. Learn.* 37, 887–901. doi: 10.1111/jcal.12538

Ma, J., and Nickerson, J. V. (2006). Hands-on, simulated, and remote laboratories: A comparative literature review. *ACM Comput. Surv.* 38, 1–24. doi: 10.1145/1132960.1132961

* Martinez, G., Naranjo, F. L., Perez, A. L., Suero, M. I., and Pardo, P. J. (2011). Comparative study of the effectiveness of three learning environments: Hyper-realistic virtual simulations, traditional schematic simulations and traditional laboratory. *Phys. Rev. Spec. Top. Phys. Educ. Res.* 7:020111. doi: 10.1103/PhysRevSTPER.7.020111

* Merkouris, A., Chorianopoulou, B., Chorianopoulos, K., and Chrissikopoulos, V. (2019). Understanding the notion of friction through gestural interaction with a remotely controlled robot. *J. Sci. Educ. Technol.* 28, 209–221. doi: 10.1007/s10956-018-9760-2

Minogue, J., and Jones, M. G. (2006). Haptics in education: Exploring an untapped sensory modality. *Rev. Educ. Res.* 76, 317–348. doi: 10.3102/00346543076003317

* Olympiou, G., and Zacharia, Z. C. (2012). Blending physical and virtual manipulatives: An effort to improve students' conceptual understanding through science laboratory experimentation. *Sci. Educ.* 96, 21–47. doi: 10.1002/sce.20463

Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *J. Educ. Stat.* 8, 157–159. doi: 10.2307/1164923

Perkins, K., Podolefsky, N., Lancaster, K., and Moore, E. (2012). Creating effective interactive tools for learning: Insights from the PhET interactive simulations project. In T. Amiel and B. Wilson (Eds.), Proceedings of EdMedia 2012–World Conference on Educational Media and Technology (pp. 436–441). Waynesville, NC: Association for the Advancement of Computing in Education. Avilable at: https://www.learntechlib.org/primary/p/40781/

* Pyatt, K., and Sims, R. (2012). Virtual and physical experimentation in inquiry-based science labs: Attitudes, performance and access. *J. Sci. Educ. Technol.* 21, 133–147. doi: 10.1007/s10956-011-9291-6

Rangarajan, K., Davis, H., and Pucher, P. H. (2020). Systematic review of virtual haptics in surgical simulation: A valid educational tool? *J. Surg. Educ.* 77, 337–347. doi: 10.1016/j.jsurg.2019.09.006

Renken, M. D., and Nunez, N. (2013). Computer simulations and clear observations do not guarantee conceptual understanding. *Learn. Instr.* 23, 10–23. doi: 10.1016/j.learninstruc.2012.08.006

Schneider, B., Krajcik, J., Lavonen, J., Salmela-Aro, K., Klager, C., Bradford, L., et al. (2022). Improving science achievement—is it possible? Evaluating the efficacy of a high school chemistry and physics project-based learning intervention. *Educ. Res.* 51, 109–121. doi: 10.3102/0013189X211067742

* Sullivan, S., Gnesdilow, D., Puntambekar, S., and Kim, J. S. (2017). Middle school students' learning of mechanics concepts through engagement in different sequences of physical and virtual experiments. *Int. J. Sci. Educ.* 39, 1573–1600. doi: 10.1080/09500693.2017.1341668

Suurmond, R., Van Rhee, H., and Hak, T. (2017). Introduction, comparison and validation of meta-essentials: A free and simple tool for meta-analysis. *Res. Synth. Methods* 8, 537–553. doi: 10.1002/jrsm.1260

Sweller, J., Van Merriënboer, J. J. G., and Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educ. Psychol. Rev.* 31, 261–292. doi: 10.1007/s10648-019-09465-5

* Tarng, W., Lee, C. Y., Lin, C. M., and Chen, W. H. (2018). Applications of virtual reality in learning the photoelectric effect of liquid crystal display. *Comput. Appl. Eng. Educ.* 26, 1956–1967. doi: 10.1002/cae.21957

Thomas, R. M. (2005). *Comparing theories of child development*, *6th Edn*. Belmont, CA: Thomson/Wadsworth.

Toth, E. E., Ludvico, L. R., and Morrow, B. L. (2014). Blended inquiry with hands-on and virtual laboratories: The role of perceptual features during knowledge construction. *Interact. Learn. Environ.* 22, 614–630. doi: 10.1080/10494820.2012.693102

* Triona, L. M., and Klahr, D. (2003). Point and click or grab and heft: Comparing the influence of physical and virtual instructional materials on elementary school students' ability to design experiments. *Cogn. Instr.* 21, 149–173. doi: 10.1207/S1532690XCI2102_02

* Wang, T. L., and Tseng, Y. K. (2018). The comparative effectiveness of physical, virtual, and virtual-physical manipulatives on third-grade students' science achievement and conceptual understanding of evaporation and condensation. *Int. J. Sci. Math. Educ.* 16, 203–219. doi: 10.1007/s10763-016-9774-2

* Winn, W., Stahr, F., Sarason, C., Fruland, R., Oppenheimer, P., and Lee, Y.-L. (2006). Learning oceanography from a computer simulation compared with direct experience at sea. *J. Res. Sci. Teach.* 43, 25–42. doi: 10.1002/tea.20097

* Yuan, Y., Lee, C. Y., and Wang, C. H. (2010). A comparison study of polyominoes explorations in a physical and virtual manipulative environment. *J. Comput. Assist. Learn.* 26, 307–316. doi: 10.1111/j.1365-2729.2010.00352.x

Zacharia, Z. C. (2015). Examining whether touch sensory feedback is necessary for science learning through experimentation: A literature review of two different lines of research across K-16. *Educ. Res. Rev.* 16, 116–137. doi: 10.1016/j.edurev.2015.10.001

* Zacharia, Z. C., and Constantinou, C. P. (2008). Comparing the influence of physical and virtual manipulatives in the context of the physics by inquiry curriculum: The case of undergraduate students' conceptual understanding of heat and temperature. *Am. J. Phys.* 76, 425–430. doi: 10.1119/1.2885059

* Zacharia, Z. C., and De Jong, T. (2014). The effects on students' conceptual understanding of electric circuits of introducing virtual manipulatives within a physical manipulatives-oriented curriculum. *Cogn. Instr.* 32, 101–158. doi: 10.1080/07370008.2014.887083

* Zacharia, Z. C., Loizou, E., and Papaevripidou, M. (2012). Is physicality an important aspect of learning through science experimentation among kindergarten students? *Early Child. Res. Q.* 27, 447–457. doi: 10.1016/j.ecresq.2012.02.004

* Zacharia, Z. C., and Olympiou, G. (2011). Physical versus virtual manipulative experimentation in physics learning. *Learn. Instr.* 21, 317–331. doi: 10.1016/j.learninstruc.2010.03.001

Zhang, L., Kirschner, P. A., Cobern, W. W., and Sweller, J. (2022). There is an evidence crisis in science educational policy. *Educ. Psychol. Rev.* 34, 1157–1176. doi: 10.1007/s10648-021-09646-1

TYPE Original Research
PUBLISHED 11 April 2023
DOI 10.3389/feduc.2023.1071759

OPEN ACCESS

EDITED BY
Wang-Kin Chiu,
The Hong Kong Polytechnic University,
China

REVIEWED BY
Sarah Pila,
Northwestern University,
United States
Anna Shvarts,
Utrecht University,
Netherlands

*CORRESPONDENCE
Stacy A. Doore
✉ sadoore@colby.edu

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to STEM Education, a section of the journal Frontiers in Education

RECEIVED 16 October 2022
ACCEPTED 08 February 2023
PUBLISHED 11 April 2023

CITATION
Doore SA, Dimmel J, Kaplan TM, Guenther BA and Giudice NA (2023) Multimodality as universality: Designing inclusive accessibility to graphical information. *Front. Educ.* 8:1071759.
doi: 10.3389/feduc.2023.1071759

Multimodality as universality: Designing inclusive accessibility to graphical information

Stacy A. Doore[1*†], Justin Dimmel[2†], Toni M. Kaplan[3], Benjamin A. Guenther[3,4] and Nicholas A. Giudice[3,5†]

[1]Immersive Navigation Systems and Inclusive Tech Ethics (INSITE) Laboratory, Department of Computer Science, Colby College, Waterville, ME, United States, [2]Research in STEM Education (RiSE) Center, School of Learning and Teaching, University of Maine, Orono, ME, United States, [3]Virtual Environment and Multimodal Interaction (VEMI) Laboratory, University of Maine, Orono, ME, United States, [4]Department of Psychology, College of Liberal Arts and Sciences, University of Maine, Orono, ME, United States, [5]Spatial Computing Program, School of Computing and Information Science, University of Maine, Orono, ME, United States

Graphical representations are ubiquitous in the learning and teaching of science, technology, engineering, and mathematics (STEM). However, these materials are often not accessible to the over 547,000 students in the United States with blindness and significant visual impairment, creating barriers to pursuing STEM educational and career pathways. Furthermore, even when such materials are made available to visually impaired students, access is likely through literalized modes (e.g., braille, verbal description), which is problematic as these approaches (1) do not directly convey spatial information and (2) are different from the graphic-based materials used by students without visual impairment. The purpose of this study was to design and evaluate a universally accessible system for communicating graphical representations in STEM classes. By combining a multisensory vibro-audio interface and an app running on consumer mobile hardware, the system is meant to work equally well for all students, irrespective of their visual status. We report the design of the experimental system and the results of an experiment where we compared learning performance with the system to traditional (visual or tactile) diagrams for sighted participants (n=20) and visually impaired participants (n= 9) respectively. While the experimental multimodal diagrammatic system (MDS) did result in significant learning gains for both groups of participants, the results also revealed no statistically significant differences in the capacity for learning from graphical information across both comparison groups. Likewise, there were no statistically significant differences in the capacity for learning from graphical information between the stimuli presented through the experimental system and the traditional (visual or tactile) diagram control conditions, across either participant group. These findings suggest that both groups were able to learn graphical information from the experimental system as well as traditional diagram presentation materials. This learning modality was supported without the need for conversion of the diagrams to make them accessible for participants who required tactile materials. The system also provided additional multisensory information for sighted participants to interpret and answer questions about the diagrams. Findings are interpreted in terms of new universal design principles for producing multisensory graphical representations that would be accessible to all learners.

KEYWORDS

blind and low vision, multimodal interface, accessible STEM graphics, multisensory interactions, learning system, universal design, assistive technologies

1. Introduction

Traditional learning materials used in mainstream science, engineering, technology, and mathematics (STEM) classrooms rely heavily on graphics and images to efficiently convey complex concepts (Kress and Van Leeuwen, 2006). However, these materials are often inaccessible to students with blindness or significant visual impairment (BVI), and this inaccessibility creates significant barriers to STEM educational and career pathways. Current statistics on the number of school-age children that meet the federal definition of visual impairment[1] (including blindness) are often difficult to obtain due to the ways in which incidence data is defined by different states (National Center for Education Statistics (NCES), 2022), and some research suggests that the federal child count underestimates the incidence of visual impairment (Schles, 2021). According to the 2021 American Community Survey (ACS), there are 7.5 million (2.5%) Americans, who are blind or have low vision, including approximately 547,000 children with severe vision difficulty under the age of 18[2] (U.S. Census Bureau, 2020). Of those students, there were approximately 55,249 United States children, youth, and adult students in educational settings who were classified as legally blind[3] (American Printing House for the Blind (APH), 2019). School success opportunities and outcomes can have lifelong impact on BVI individuals. Of the nearly 4 million civilian non-institutionalized working age adults (18–64) with a visual impairment, only 2 million (50%) are employed, another 250,000 (5%) working adults with a visual impairment are classified as unemployed (but still looking for work), with the remaining 1.8 million (45%) of adults with visual impairments classified as not actively engaged in the labor force (McDonnall and Sui, 2019; U.S. Census Bureau, 2020). This compares to 136 million (77%) civilian non-institutionalized working age adults (18–64) without a disability who are employed, 8 million (5%) unemployed (but still looking for work), with the remaining 33 million (18%) of the adult population without a disability classified as not actively engaged in the labor force (U.S. Census Bureau, 2020).

This large disparity between the employment rates for BVI adults and the general population without a disability (50% vs. 77%) helps to motivate our work to improve information access for advancing into STEM related careers and the opportunities that are available with advanced STEM education. As workplaces become more automated, future labor market skills needed to maintain United States progress and innovation will require more diversity of perspectives for complex problem solving, therefore "all learners must have an equitable opportunity to acquire foundational STEM knowledge" (Honey et al., 2020). In order to support these equitable opportunities in STEM, there is a profound need for accessible STEM training tools and learning materials to provide learning access across future labor contexts and for people of all ages. NSF STEM participation data does not breakdown participation by disability type (Blaser and Ladner, 2020). What is well documented is that the dearth of accessible materials for BVI learners at all levels and how this presents acute challenges for inclusive STEM courses (Moon et al., 2012). To take one example, the study of geometry, as manifested in secondary schools, is inextricably bound with what has been described as the *diagrammatic register* – a communication modality in which mathematical concepts are conveyed through logical statements (written in words) that are linked to diagrams (Dimmel and Herbst, 2015). The primary challenge for BVI learners is that geometry diagrams visually convey properties that do not explicitly describe spatial information in the accompanying text, such as whether a point is on a line, or whether two lines intersect. Thus, verbal descriptions containing additional information are necessary to make the diagrams accessible to BVI learners, however, these longer descriptions use additional words that can increase the cognitive load of making sense of the representation, with the long descriptions often still failing to convey key spatial content (Doore et al., 2021). As a result, BVI learners spend significantly more time and are far less accurate than their sighted peers in interpreting diagrammatic representations due to the lack of consistent standards for graphical content metadata, including description annotations (Sharif et al., 2021; Zhang et al., 2022). While extended length description recommendations for graphical representations have evolved and improved over time (Hasty et al., 2011; W3C, 2019), few guidelines for natural language descriptions of diagrams, charts, graphs, and maps are grounded in any theoretical framework with some notable exceptions using category theory (Vickers et al., 2012) spatial cognition theory (Trickett and Trafton, 2006), semiotic theory (Chandler, 2007), and linguistic theory (Lundgard and Satyanarayan, 2021). We view our work as complementary to this body of theoretically grounded research, embedding structured natural language descriptions into accessible multisensory data representations that use haptics, spatial audio, and high contrast visuals to help with the interpretation of graphic information.

Beyond the challenges of creating accessible information ecosystems in classrooms for all learners, the STEM visualization access challenge has received growing attention at a broader societal level as the use of graphical representations has been shown to play an important role in conveying abstract concepts and facilitating the deeper meaning of scientific texts (Khine, 2013). The information access gap inevitably contributes to the lower academic performance observed in math and science among BVI students in comparison both to other subjects and also to their non-visually impaired peers within STEM disciplines (Cryer et al., 2013). The limited availability of blind-accessible materials can also force teachers to adopt content that employs phrasing, structure, or terminology that does not correspond with the teacher's preferred method of instruction or intended curriculum. This lack of access to educational materials can

1 Federal regulations define visual impairment (including blindness) as "an impairment in vision that, even with correction, adversely affects a child's educational performance" [34CFR Sec. 300.8(c)(13)]. Some states have elaborated on this definition by specifying minimum levels of visual acuity or a restriction in the visual field. Thus, a child may qualify as having a visual impairment in one state but may not qualify in another https://ies.ed.gov/ncser/pubs/20083007/index.asp.

2 The children referred to range in age from 0 to 17 years and only included those children that had serious difficulty seeing even when wearing glasses as well as those that are blind.

3 The students referred to range in age from 0 to 21 years as well as certain qualifying adult students and only included those students with vision loss that functioned at/met the legal definition of blindness. Legal blindness is a level of vision loss that has been defined by law to determine eligibility for benefits. It refers to explicitly to those who have a central visual acuity of 20/200 or less in the better eye with the best possible correction, or a visual field of 20 degrees or less.

make classroom learning and information interpretation difficult, resulting in BVI students falling significantly behind standard grade level content (Lundgard and Satyanarayan, 2021).

There is thus an urgent need for a universal accessibility solution providing inclusive information access to STEM content supporting the same level of learning, understanding, and representation—i.e., functionally equivalent performance—for all learners. By universal, we mean the solution should use only those accessibility supports that could reasonably be expected to be familiar and available to all learners—i.e., the solution would not require specialized hardware or knowledge of specialized systems of communication, such as braille (National Federation of the Blind (NFB), 2009). By *functionally equivalent,* we mean the representations built up from different modalities will be associated with similar behavioral performance on STEM tasks (e.g., accuracy and success rate; Giudice et al., 2011). Evidence for such functional equivalence has been observed across many tasks and is explained by the development of a sensory-independent, 'spatial' representation in the brain, called the spatial image, which supports similar (i.e., statistically equivalent) behavior, independent of the learning modality (for reviews, see Loomis et al., 2013; Giudice, 2018). Functional equivalence has been demonstrated with learning from many combinations of inputs (visual, haptic, spatialized audio, spatial language), showing highly similar behavioral performance across a range of inputs and spatial abilities including spatial updating (Avraamides et al., 2004), target localization (Klatzky et al., 2003), map learning (Giudice et al., 2011) and forming spatial images in working-memory (Giudice et al., 2013).

The purpose of the current study was to investigate how effectively working age adults could learn graphical-based STEM content information from a universally designed interface that was developed to support functional equivalence across visual and non-visual modalities for representing diagrams. We asked: How effectively do multisensory inputs (high contrast visuals, spatial language, and haptics) convey functionally equivalent spatial information for learning concepts that are represented in diagrams? We investigated this question by developing and testing a multisensory diagram system that was designed to be accessible to all learners.

2. Background

2.1. Universal design for assistive technologies

Our focus on all learners was motivated by two considerations. One, BVI learners face significant social challenges in school, where impromptu group discussion and peer-to-peer learning are important components of social and behavioral skill development (Smith et al., 2009). Inclusive classrooms are increasingly the most common educational settings among BVI students, with over 80% of this demographic attending local public schools and spending most of their time in inclusive classrooms alongside of their sighted peers (Heward, 2003; American Printing House for the Blind (APH), 2019). Two, a universal design approach is thus advantageous because it reduces barriers for BVI learners to participate in peer-mediated classroom activities (e.g., group work) – when everyone is using the same resources, there is no reason for the BVI learners to receive special accommodation. This is also an important consideration as whole class discussions that occur naturally in inclusive settings play a crucial role in the development of social, linguistic, and behavioral skills, as well as improve conceptual understanding and overall academic performance (Smith et al., 2009; Voltz et al., 2016).

2.2. Multisensory learning

Apprehending information through multiple sensory modalities is beneficial for everyone, not only those for whom a sensory accommodation was initially designed (e.g., Yelland, 2018; Abrahamson et al., 2019). In the 20 years since Mayer's seminal paper "Multimedia Learning" (Mayer, 2002), hundreds of studies have investigated how complementary sensory modalities, such as pictures and text, can enhance the acquisition and retention of information. How closed captioning has been adopted and integrated into educational, professional, and recreational videos is one example. Closed-captioning benefits deaf and non-hard of hearing viewers alike (Kent et al., 2018; Tipton, 2021). The availability of closed captioning across media reflects not only a commitment to accessibility but also provides empirical examples, at scale, that illustrate the redundancy principle.

The redundancy principle hypothesizes that simultaneous presentations of the same information *via* different modes allows modality-independent sensory processing to occur simultaneously: Two cognitive systems can process the same information in parallel (Moreno and Mayer, 2002). Reading closed captions taxes visual working memory, while hearing spoken words taxes auditory working memory. These processes are independent, which means reading captions while simultaneously listening to spoken words allows for both the visual and auditory systems to work synergistically toward apprehending the information that is represented in written (visual) and spoken (auditory) words. Redundancy, when partitioned across independent sensory modalities, helps learners build and retain conceptual (i.e., mental) models (Moreno and Mayer, 2002), which are integrated representations of spatial information about objects and relations.

2.3. Spatial mental models

Model theory asserts that people translate a perceived spatial configuration into a mental model and then use this mental representation to problem-solve and make inferences on spatial information (Johnson-Laird, 1983, 2010; Johnson-Laird and Byrne, 1991). Under the best of conditions, spatial reasoning problems are difficult to solve using language alone (Ragni and Knauff, 2013). For example, describing something as simple as how to locate the reception desk within a hotel lobby is both a complex description task (for the person doing the describing) and a difficult non-visual navigation task (for the BVI person who needs directions to navigate) because there are no tools (e.g., a standardized coordinate system) for providing spatial references within the lobby (or other similar indoor environments). As such, the non-visual navigation task for solving what we call the "lobby problem," i.e., independently finding the check-in desk from a hotel's main entrance, or the elevator from the check-in desk, or the hotel restaurant down a long hallway from the lobby can be extremely challenging. Instructional graphics present a similar challenge where the typical accessibility solution is a poorly structured (and all too often ambiguous) description from a teacher or instructional aide. The adage that 'a picture is worth a thousand words' is most certainly true in that humans can process complex visual information in an image to understand spatial configurations, relationships, and be able to make inferences on their meaning far more quickly and efficiently than it would take to verbally describe a complex graphic.

There are several factors that influence spatial information processing using language to form mental models, such as the number of required models to solve a reasoning problem (Johnson-Laird, 2006), presentation order (Ehrlich and Johnson-Laird, 1982), use of transitive/non-transitive relations (Knauff and Ragni, 2011), binary/*n*-place relations (Goodwin and Johnson-Laird, 2005), and the differences in spatial reasoning on determinate/indeterminate problems (Byrne and Johnson-Laird, 1989). In many cases, sighted annotators (and in turn automated image captioning systems) often use non-transitive spatial relations such as "next to" or "contact" or "on the side" instead of transitive relations such as "left of" or "in front of" (Knauff and Ragni, 2011) to describe the spatial arrangements in images. Imagine the difficulties BVI students would face if they had to reason about a 100-point scatterplot that they could not see and instead, were provided with a list of 100 ordered pairs accompanied by a set of vague descriptions about their relative spatial positions (e.g., "point A (3,6) is on the side of point K (3,7) which is below point G"). Instead, we argue that spatial information must be explicitly incorporated into accessible learning systems to reflect current multisensory learning and model theory. This study investigates how high contrast visuals, sonification, vibrotactile haptic feedback, and spatial information descriptions collectively affect information retention, when compared to traditional accessibility solutions.

3. Design and development

3.1. The multimodal diagram system

We designed and developed a multimodal diagram system (MDS) to investigate how effectively one platform could provide multisensory representations that would be accessible to all learners, especially those who have visual impairments. The MDS was specifically designed to be widely accessible and practical for diverse user populations, such as the broad spectrum of users with vision impairments and sighted users who require increased multimodal information access. The prototype system was designed on the iOS platform to leverage the many embedded universal design features in the native Apple iPhone UI, which accounts for why the vast majority of BVI smartphone users (72–80%) prefer to use iOS-based devices (Griffin-Shirley et al., 2017; WebAim.org, 2019). The MDS has two components: (1) the MDS *vibro-audio interface* mobile application and (2) an associated website that hosts a diagram library and an online diagram annotation/authoring tool for use by diagram creators.

3.1.1. Vibro-audio interface mobile application

The MDS renders a high-contrast diagram on screen and provides audio and haptic feedback when the screen is touched, making diagram information access possible with or without vision. The MDS was designed to be used *via* single finger screen scanning. A short vibration is triggered when the user moves their finger over the bounds of an onscreen element (i.e., moving from one element to another, such as from the front paneling of a house to the door). They can move their hand to follow/trace the vibratory element or listen to hear its name (tapping will repeat the auditory label). The dimensions of on-screen elements were informed by prior research into vibro-haptic interface design. For the MDS, the minimum width of lines was 4 mm and the minimum gap between lines was 4 mm (Palani et al., 2020). The vibration feature at the edge of an element was designed to be analogous to the raised lines between features of a traditional embossed tactile diagram and was implemented using the default iOS notification vibration.

Haptic/audio redundancy was integrated with the design: While a user's finger touches a graphic element, such as the red circle in Figure 1, a constant element specific audio background tone is played. In addition, the name of the element is read *via* text-to-speech, and after a brief pause, a description of the element is read if the user's finger remains within the red circle element. If the user moves their finger to enter a new element, such as the purple section in Figure 1, a vibration is triggered as their finger changes elements, then the unique audio tone, label and description begins again for this new graphic element. This procedure is based on guidance from earlier multimodal research (Choi and Walker, 2010).

3.1.2. Multimodal diagrammatic system interface information flow

The MDS conveyed graphical information through images (visual), spoken words (auditory), sounds (auditory), and vibrations (haptic), where there were redundancies among the visual, auditory, haptic modalities along with kinesthetic cues (e.g., hand-movement and gestures). The MDS visually represented points as high-contrast, color-filled circles/vertices, lines/curves as 4 mm width high-contrast, colored line segments, and shapes/regions as high-contrast, color-filled areas. Simultaneously, points, lines, curves, and regions were each represented as distinct audio tones. The tones were programmatically generated for each element in a diagram by incrementally shifting a 180 Hz sine wave tone up in pitch depending on the number of elements in the diagram to assure that each element was represented by a distinguishing tone. Also simultaneously, points/lines/curves were represented haptically through the phones vibration motor that was activated whenever a finger touched that x-y point on the screen.

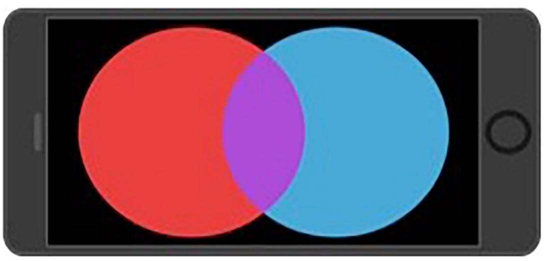

FIGURE 1
A Venn diagram using red and blue circles overlapping in a purple section, rendered here as it might be visually presented using the MDS.

When a diagram is loaded *via* the MDS, the system uses text-to-speech to read (auditory) the diagram title and, if present, an instructors note (e.g., instructing the user to begin their exploration at the bottom of the diagram). Element labels (e.g., "point p," "line l") were provided through native iOS text-to-speech and were played whenever a user's finger entered the bounds of an element. Text-to-speech element descriptions followed 1 second (s) after the element label was read if the user remained within that element.

4. Materials and methods

The user study employed a perception-based (rather than a memory-based) information access task, where participants had access to the diagrams while they simultaneously completed worksheets related to the content. To control for pre-existing knowledge (i.e., variability in pre-test scores), pre-test and active-test worksheets were used to calculate normalized information gain scores. A finding of similar information gain between the MDS interface and traditional hardcopy stimuli would indicate that the MDS system is equivalently effective in conveying non-visual information. This design was motivated by previous work in the education and educational gaming literature (Furió et al., 2013). Similar procedures are typically used in education technology research to provide a "consistent analysis over diverse student populations with widely varying initial knowledge states" (Hake, 1998).

Users completed pre-test questions to establish their pre-existing knowledge on the diagram content. They then completed diagram content related worksheets while using diagrams in two different modal conditions: (1) using a traditional diagram (a visual or embossed/tactile diagram, between sighted and blind users, respectively), or (2) using the experimental MDS interface. The test worksheets used in this experiment were identical in format but not content to those used in the pre-test and were designed to emulate worksheets employed in standard STEM curricula. As the evaluation was designed as a perceptual task to determine whether the MDS could provide access for learning new information, not how well it could facilitate recall or mental representation of the information, the active test worksheet was completed with simultaneous access to the diagram. This pre-test/active test design was used so that normalized information gain could be calculated for each diagram.

4.1. Hypothesis

We hypothesized that the MDS interface would provide a *functionally equivalent* information access solution, resulting in similar results for worksheet accuracy and time to completion between worksheets completed with diagram access using a control condition: (1) traditional tactile stimuli (BVI control condition), or (2) visually-presented stimuli (sighted control condition). That is, we postulated that the use of the multisensory interface would allow participants with BVI to function at equivalent levels to their sighted peers.

> *H0*: There will be no significant difference between groups (BVI, Sighted) completing worksheets in each condition (MDS, Control) providing a *functionally equivalent* information access experience.

This hypothesis is based on pilot user testing and previous work on functional equivalency (Giudice et al., 2013) that demonstrates the efficacy of vibro-audio interfaces in facilitating nonvisual access to spatial information using multimodal maps (Brock et al., 2010) touchscreen haptics (Palani and Giudice, 2016); spatial tactile feedback (Yatani et al., 2012) and related research on the application of multimodal interactive tools in education (Cairncross and Mannion, 2001; Moreno and Mayer, 2007). Previous work evaluating a vibro-audio interface noted slower encoding *via* learning with this type of interface in comparison to visual and traditional tactile graphics for both sighted and BVI users, although behavioral performance on testing did not differ (Giudice et al., 2012). Despite this difference in learning time, the same study found that the overall learning and mental representation of the diagram information (e.g., graphs, figures, and oriented polygons) was not reliably different between types of presentation modalities. Based on previous studies, we anticipated that with increased geometric complexity of the experimental diagrams, there could be increased worksheet completion times when using the MDS experimental interface. Similar testing of vibro-audio interfaces has found that nonvisual tracing of lines (audio or vibrotactile) rendered on a flat surface (e.g., touch screen) can be more challenging than following lines visually or using embossed tactile graphics (Giudice et al., 2012).

While this earlier work dealt with different STEM application domains, e.g., diagrams, shapes, maps, it was critical in the development and evaluation of this new multisensory interface in: (1) determining what parameters led to the most perceptually salient stimuli, (2) showing that using these multisensory stimuli led to accurate learning, mental representations, and other cognitive tasks using the interface, and (3) that it could support similar learning as was possible from existing/established modes of nonvisual information access (i.e., hardcopy tactile renderings). In other words, the early work dealt with design optimization and determining efficacy (e.g., does this system work or can stimulus x be learned using this approach?) By contrast, the pedagogy and motivation in this study is different, as we are now explicitly studying the nature of the learning and comparing this multisensory approach to existing *de facto* approaches using touch or vision between sighted and blind groups. Without this previous work, it would not be possible to use this interface here with any *a priori* knowledge of its efficacy. Our comparisons in this paper extend the previous work in multiple ways: (1) we are using very different STEM stimuli, (2) assessing its use in a knowledge gain task, and (3) comparing its use with both sighted and blind individuals (and their respective controls). An additional unique contribution of this study is the focus on the UDL nature of the system. Not only is the system being considered as an accessibility support for blind people but is conceptualized (and evaluated) as a universal support for all potential users, which has important applications for the use of multisensory devices for supporting generalized learning in a variety of STEM educational and vocational settings.

4.2. Methods

4.2.1. Participants

The study included 29 working age adult participants: 20 participants (20) without vision impairment and nine participants

(9) with legally defined blindness. We were able to recruit a reasonably matched sample of BVI and sighted participants across age, gender, and education (Tables 1, 2). As age and gender are not critical factors in the outcomes of this study, we only report these participant data in the aggregate. This initial experiment recruited working age adults evaluate the MDS efficacy across a broad age range of adult users that would be representative of a variety of demographic groups (e.g., college and vocational learners). Recruitment was conducted through direct contact with people who have previously participated in lab studies, via a study recruitment ad distributed on several listservs for blind and vision loss communities, and by posting a bulletin board study recruitment ad in the community grocery store near the University.

The unbalanced design across participants reflects the typical challenges of recruiting research participants that are visually impaired. However, this sample size was sufficiently powered and is a similar size of traditional usability studies aimed at assessing the efficacy of assistive technology interface/device functionality for BVI populations (Schneiderman et al., 2018). The studies were reviewed and approved by the university's Institutional Review Board and all participants provided their written informed consent to participate in this study. All participants self-reported as having at least some college with several participants in both groups reporting they held graduate degrees. All participants in both groups reported as daily smartphone users. All BVI participants reported as being exclusively iPhone users, which is consistent with previous research on smartphone platform preference in the BVI community (WebAim.org, 2019). Among sighted participants smart phone usage was reported at 30% Android, 70% iPhone.

4.2.2. Test interface

All participants used the iPhone-based experimental non-visual interface in the default iOS accessibility mode with a screen curtain on, thus completely disabling the phone's visual display. This was implemented to prevent any possible visual access to the presented diagram.

4.2.3. Test science, technology, engineering, and mathematics content

Diagrams in all conditions were designed to provide equivalent information and be as similar as possible, while still representing graphical rendering typical of the given modality. These specific diagrams were selected to represent topics normally presented graphically in a STEM curriculum. The two diagrams selected for use in this study were (1) layers of the atmosphere and (2) a helium atom (Figure 2). The images were created in a commercially available presentation slide platform and were based on diagrams of the same subjects from the American Printing House (APH) for the Basic Science Tactile Graphics set (American Printing House for the Blind (APH), 1997), which were used as a benchmark for BVI participants. The traditional visual diagrams used in the study were also based off the APH kit examples and adapted to resemble standard colored visual diagrams with text-based labels and a description key. Response protocol worksheets were designed to incorporate questions that demonstrated the efficacy of the interface in presenting both descriptive and spatial information. For example, each worksheet included questions regarding size and/or relative location of diagram elements, in addition to content questions regarding descriptions or functionality of the elements.

TABLE 1 Participant demographics summary.

Group	*n*=()	Sex	Age Range/ Mean	Highest education level
Participants w/ legally defined blindness	9	6 F, 3 M	21–70 M = 49.3	4 some college
				4 undergrad degree
				1 graduate degree
Participants w/o vision impairment	20	8 F, 12 M	19–35 M = 22.6	15 some college
				3 undergrad degree
				2 graduate degree

4.2.4. Test procedure

A within-subjects, mixed factorial design was used in the experiment. Within-subjects factors were diagram type and presentation mode, with visual status being a between-subjects factor. All participants completed two pre-study worksheets based on the diagram content used in the study. Participants were then given access to one of the diagrams in each modal condition. The modal conditions differed slightly depending on participants' visual status. All participants completed a common condition using the experimental MDS interface, while their control/benchmark condition varied, with BVI participants using a hardcopy vacuum-formed tactile diagram (the gold standard for tactile-based renderings) and sighted participants using a visual diagram as their control/benchmark. Importantly for such cross-modal comparisons, the diagram elements were identical in each condition and the diagrams were scaled to the same size across condition. Diagram labels were provided using text-to- speech in the experimental MDS diagram condition (e.g., *Protons are positively charged particles in the nucleus of the atom*), using written text in the visual diagram condition, and given verbally by an experimenter acting as a learning assistant in the hardcopy tactile diagram condition.

The two conditions were administered in three phases: a pre-test phase, a practice phase, and an active-test phase, which was followed by a post-study questionnaire. Condition order and diagram presentation were counterbalanced to avoid order effects. In the pre-test phase, participants completed baseline-knowledge pre-tests for both diagrams. The pre-test percent accuracy score was used to represent participants' *a priori* knowledge of the content. Worksheets were given one at a time, and participant completion time was logged for each sheet. Worksheets were scored based on the number of questions answered correctly, out of the total number of questions (i.e., percent accuracy). This was then used to calculate normalized information gain for each participant. To ensure participants were able to demonstrate information gain, only those who initially made two or more errors were deemed eligible to continue in the study. Four sighted participants completed the

TABLE 2 BVI Participant demographics.

Participant number	Etiology of blindness	Onset age	Residual vision	Highest education level	Diagram use frequency
P1	Lebers congenital amaurosis	Birth	Light/ dark perception	Undergrad. degree	None
P2	Lebers congenital amaurosis	Birth	Light/dark perception	Graduate degree	Monthly
P3	Pathological myopia	45	Light/dark perception in right eye, Fuzzy colors	Undergrad. degree	None
P4	Retinitis pigmentosa, atypical, with cone dystrophy	25	Light/dark perception, some functional peripheral	Undergrad. degree	Weekly
P5	Retinitis pigmentosa	11	Light/dark perception	Some college	None
P6	Retinopathy of prematurity	Birth	Light/dark perception	Undergrad. degree	Yearly
P7	Glaucoma	16	Light/dark perception	Some college	Yearly
P8	Unknown	17	Light/dark perception	Some college	None
P9	Congenital cataracts, glaucoma	50	Light/dark perception	Some college	Monthly

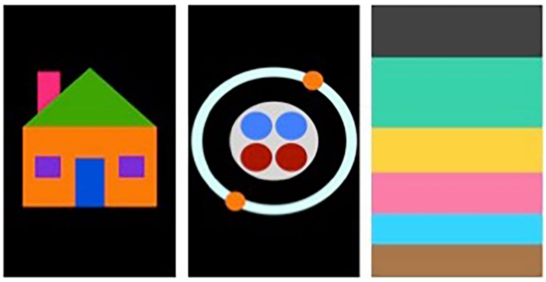

FIGURE 2
Example visualizations of the audio-haptic diagrams from left to right: practice diagram (house), atom diagram, and atmosphere diagram. Each color represents a unique element in the diagram. The phone screen was disabled *via* screen curtain throughout the study, so participants did not have visual access to the diagrams.

"Atom" diagram pre-test worksheet without any errors (i.e., earning a ceiling score), therefore, they did not continue the study and were replaced with new sighted participants. As we were looking for pre-test/active-test differences, it would not be possible to measure these differences if a participant's pre-test was already at the ceiling, making the active-test data (if included) irrelevant.

In the practice phase, participants were provided an opportunity to practice with using the house diagram before each condition, either practicing with the experimental MDS interface (experimental condition) or traditional (tactile or visual) diagram (control conditions). All participants were asked simple spatial configuration questions about the sample diagrams, (e.g., for the house sample, *How many windows are present? Where are they relative to the door? What side of the house is the chimney on?*). These questions served as a criterion test to ensure all participants achieved basic competency using the interface before moving on to the experimental trials and all participants were able to answer these questions during the practice phase. During the active-test phase, participants completed a worksheet with access to the test diagram. Worksheet completion time was measured as the duration of time required for the participants to complete the worksheets.

5. Results

5.1. Confirmation of learning

Analyses of descriptive statistics were conducted to compare prior (baseline) knowledge (based on pre-worksheet accuracy) and to confirm the presence of learning (by comparing pre- and active-worksheet accuracy) for BVI and sighted participants. These analyses were conducted by collapsing across diagram (atom vs. atmosphere) and mode (MDS vs. control). Pre-Accuracy descriptive results suggest the two groups were significantly different in their prior knowledge of using diagrams evident from the pre-worksheet accuracy mean percentage and the variability represented by the range of scores (Table 3).

Comparing data for each measure in the MDS and control conditions revealed remarkably similar values. For instance, there was only a 4% difference in active-worksheet accuracy between MDS (92%) and control (96%) conditions ($t(27) = 1.3$, $p = 0.21$). This interaction was examined *via Post-hoc* comparisons of pre-and active-worksheet accuracy for BVI and sighted participants. Independent samples t-tests revealed prior knowledge (based on pre-worksheet accuracy) for BVI participants (13.7%) was significantly less than that of sighted participants (31.2%) $t(27) = -4.16$, $p < 0.01$.

In other words, the lower overall accuracy in BVI participants was due to significantly lower pre-test (but not active-test) accuracy when compared to sighted participants. Importantly,

TABLE 3 Pre/post accuracy collapsed across diagram and mode.

Visual status	Pre-accuracy mean (%)	SD	Range	95% CI	Active-accuracy mean (%)	SD	Range	95% CI
BVI	13.7	18.5	0–50	4.5–22.8	96.1	6.6	83.3–100	92.8–99.4
Sighted	31.2	34.9	0–75	20.1–42.4	92.0	10.7	66.7–100	88.5–95.4

although each group started with different levels of prior knowledge, their final scores (as measured by the active-test worksheet) were remarkably similar.

Furthermore, weighted mean accuracy data were submitted to a 2 × 2 ((learning: pre- vs. active-worksheet accuracy) X (group: BVI vs. sighted)) mixed-model ANOVA. Learning across all participants was evident in the difference between pre-test worksheet accuracy (25%) and active-test worksheet accuracy (94%), $F(1,27) = 597.1$, η_p^2 (partial eta^2) = 0.9, $p < 0.01$. Collapsing across pre-test and active-test worksheet performance, overall accuracy was reliably lower for BVI participants (56%) than for sighted participants (64%), $F(1,27) = 7.2$, $\eta_p^2 = 0.2$, $p < 0.01$. However, this difference was likely driven by the interaction between participant group and learning mode, $F(1,27) = 19.0$, $\eta_p^2 = 0.4$, $p < 0.01$.

5.2. Information gain and completion time

In addition to the dependent variables of pre-test and active-test worksheet accuracy, the effect of the MDS interface and traditional hardcopy diagrams were also evaluated on two measures calculated to control for variance in pre-test knowledge (information gain and worksheet completion time). Individual normalized information gain scores reflect the improvement from pre- to post-test divided by the total amount of improvement possible ([gain = %posttest = %pretest]/[100-%pretest]) and were calculated for each participant's performance in both modal conditions (Hake, 1998).

Worksheet completion time was calculated by dividing the time to complete the worksheet by the number of questions participants needed to answer on the worksheet (this varied based on pre-worksheet performance). Participants completed worksheets (two in total) using both modalities (MDS and control) and both diagrams (atom and atmosphere). Each diagram could only be tested once per participant; therefore, a full 2 (mode) × 2 (diagram) within-subjects design was not possible. Therefore, the effect of mode and diagram were each considered separately (collapsing across the other factor).

5.3. Information presentation mode

Pre-Accuracy descriptive results suggest the two groups were significantly different in their prior knowledge of using diagrams evident from the pre-worksheet accuracy mean percentage and the variability represented by the range of scores (Table 4). Analyses of descriptive statistics were conducted comparing data for each measure in the MDS and control conditions. These revealed remarkably similar values. There were greater Mean information gains for both groups. Additionally, the average time spent per question to complete the worksheet was also quite similar between MDS and control conditions.

The effect of mode was evaluated *via* 2 × 2 ((mode: MDS vs. control) x (group: BVI vs. sighted)) mixed MANOVA with active-worksheet accuracy, information gain, and worksheet completion time serving as the dependent measures. Neither the multivariate main effects nor the multivariate interaction reached significance (all p's > 0.05).

5.4. Diagram type

Analyses of descriptive statistics were conducted comparing diagram type by condition and these also revealed remarkably similar values (Table 5). Again, there was only a small difference in active-worksheet accuracy in the BVI group between MDS (94%) and control (98%) conditions and in the sighted group between the MDS (90%) and control (94%) conditions. Additionally, the average time spent per question to complete the active worksheet was also quite similar across diagram types between MDS and control conditions.

The effect of diagram type was evaluated *via* a 2 × 2 ((diagram: atom vs. atmosphere) X (group: BVI vs. sighted)) mixed MANOVA with pre-worksheet accuracy, active-worksheet accuracy, gain, and worksheet completion time serving as the dependent measures. Analyses revealed significant multivariate main effects of group, Wilks' $\lambda = 0.5$, $F(4,24) = 4.3$, $\eta_p^2 = 0.4$, $p < 0.01$ and diagram, Wilks' $\lambda = 0.2$, $F(4,24) = 25.1$, $\eta_p^2 = 0.8$, $p < 0.01$, as well as a significant multivariate interaction between the two factors, Wilks' $\lambda = 0.5$, $F(4,24) = 6.2$, $\eta_p^2 = 0.5$, $p < 0.01$. Given these results, univariate main effects and interactions are presented below.

5.4.1. Pre-worksheet accuracy

There was a significant difference in pre-worksheet accuracy between BVI (14%) and sighted (31%) participants with a greater variance in pre-test accuracy observed in the sighted participants, $F(1,27) = 16.0$, $\eta_p^2 = 0.3$, $p < 0.01$ (Table 6). There was also a significant effect of diagram, $F(1,27) = 95.3$, $\eta_p^2 = 0.8$, $p < 0.01$ with greater pre-worksheet accuracy for the atom diagram (43%) than the atmosphere (2%) diagram. The interaction between diagram type and group also reached significance, $F(1,27) = 25.5$, $\eta_p^2 = 0.5$, $p < 0.01$ (Table 7).

Pre-test worksheet accuracy on the atom diagram was lower in BVI participants (24%) than that of the sighted participants (63%) with similar variability (Table 7). Independent sample t-tests revealed that was significantly less $t(27) = -4.6$, $p < 0.01$. However, there were no reliable differences between BVI (4%) and sighted (0%) participants for the atmosphere diagram, t(8.000) = 1.5, $p = 0.17$ (corrected values reported due to heterogeneity of variance, $F = 41.7$, $p < 0.01$). Additionally, paired-sample t-tests revealed that BVI participants, $t(8) = 2.9$, $p < 0.05$, and sighted participants, $t(19) = 13.3$, $p < 0.01$ had

10.3389/feduc.2023.1071759

TABLE 4 Pre/post accuracy collapsed across condition mode.

Visual status	Diagram	Gain mean (%)	SD	Range	95% CI	Pre-accur mean (%)	SD	Range	95% CI	Active accur. mean (%)	SD	Range	95% CI	Comp time mean (s)	SD	Range	95% CI
BVI	Atom	95.4	7.0	83.0–100.0	90.1–100.8	23.6	21.1	0.0–50.0	7.4–39.7	95.8	6.3	87.5–100	91.0–100.6	46.0	18.5	16.5–80.5	31.8–60.2
	Sky	96.2	7.5	83.0–100.0	90.5–102.0	3.7	7.35	0.0–16.7	−1.9-9.4	96.3	7.4	83.3–100	90.6–101.9	41.6	13.2	27.8–63.4	31.4–51.7
Sighted	Atom	79.2	25.2	33.0–100.0	67.4–90.9	62.5	21.1	0.0–75	52.6–72.4	93.1	8.6	75–100	89.1–97.1	65.7	28.8	35.0–158.0	52.2–79.2
	Sky	90.8	12.6	67.0–100.0	84.9–96.7	0 0.0	0.0	0.0–0.0	0.0–0.0	90.8	12.7	66.7–100	84.9–96.8	31.9	8.12	20.3–50.8	28.1–35.7

TABLE 5 Pre/post accuracy collapsed across diagram.

Visual status	Condition	Gain mean (%)				Pre-accur				Active accur				Comp time			
			SD	Range	95% CI	Mean (%)	SD	Range	95% CI	Mean (%)	SD	Range	95% CI	Mean (s)	SD	Range	95% CI
BVI	MDS	93.6	7.8	83.0–100.0	87.5–99.6	6.0	9.6	0.0–25.0	−1.3-13.4	94.0	7.3	83.3–100.0	88.4–99.6	38.2	12.0	16.5–52.0	29.0–47.4
	Control	98.1	5.7	83.0–100.0	93.8–100.0	21.3	22.4	0.0–50.0	4.1–38.5	98.1	5.6	83.3–100.0	93.9–102.4	49.4	17.7	31.5–80.5	35.8–62.9
Sighted	MDS	85.8	17.1	50.0–100.0	77.8–93.8	27.5	34.1	0.0–75.0	11.6–43.4	90.2	11.5	66.7–100.0	84.8–95.6	51.6	30.1	21.7–158.0	37.5–65.6
	Control	84.2	23.9	33.0–100.0	73.0–95.3	35.0	36.2	0.0–75.0	18.1–51.9	93.8	9.9	66.7–100.0	89.1–98.4	46.1	24.0	20.3–101.0	34.8–57.3

better pre-worksheet accuracy for the atom diagram as compared to the sky/atmosphere diagram.

5.4.2. Active-worksheet accuracy and information gain

For active-worksheet accuracy (Table 8), results suggest that with access to the MDS, the active worksheet accuracy for both groups were similar. In other words, neither of the main effects of group, $F(1,27) = 1.7$, $\eta_p^2 = 0.1$, $p = 0.2$, nor diagram, $F(1,27) = 0.1$, $\eta_p^2 = 0.01$, $p = 0.7$, nor the interaction, $F(1,27) = 0.3$, $\eta_p^2 = 0.01$, $p = 0.6$ reached significance.

For information gain, neither of the main effects of group $F(1,27) = 3.7$, $\eta_p^2 = 0.1$, $p = 0.1$, nor diagram $F(1,27) = 2.4$, $\eta_p^2 = 0.1$, $p = 0.1$, nor the interaction $F(1,27) = 1.9$, $\eta_p^2 = 0.1$, $p = 0.2$ reached significance.

5.4.3. Worksheet completion time

There were small differences between groups in completing the worksheet tasks, however, the main effect of group on worksheet completion time, $F(1,27) = 0.7$, $\eta_p^2 = 0.02$, $p = 0.4$ did not reach significance. The BVI participants Mean time in seconds (46 s) for the atom diagram was faster than the sighted participants Mean time (66 s) (Table 7) with the main effect of diagram $F(1,27) = 14.7$, $\eta_p^2 = 0.3$, $p < 0.01$ reaching significance. This finding is not surprising given our *a priori* prediction that more geometrically complex diagrams would result in slower non-visual diagram access with the MDS.

The interaction between diagram and participant group (see Table 4) was also significant, $F(1,27) = 8.7$, $\eta_p^2 = 0.2$, $p < 0.01$. Independent samples *t*-tests revealed worksheet completion time for the sky/atmosphere diagram was significantly longer for BVI (42 s) participants than for sighted (32 s) participants (Table 7), $t(27) = 2.4$, $p < 0.05$; however, there were no differences in worksheet completion time between groups for the atom diagram, $t(27) = -1.9$, $p = 0.07$. Paired-samples *t*-tests did not reveal a significant difference in worksheet completion time for BVI participants, $t(8) = 0.6$, $p = 0.6$. However, sighted participants took significantly longer for the atom (66 s) compared to the atmosphere (32 s) diagram (Table 7), $t(19) = 5.8$, $p < 0.01$.

The differences in the completion time results may be attributed to the lack of familiarity with non-visual learning among the sighted participants, as well as the use of inefficient tactile scanning strategies by people who are not accustomed to learning through this modality. This interpretation is consistent with other studies showing that differences in tactile scanning strategies can impact the efficiency and accuracy of information acquisition and participant performance on spatial search tasks (Ungar et al., 1996).

5.4.4. Visual status

A sub-analysis of the descriptive statistics for the BVI participant data was conducted to investigate the potential impact of any residual vision on the pre/post worksheet accuracy and completion time (Table 9). Looking at the raw data for performance of BVI participants, there does not appear to be any noteworthy differences based on visual status. Participants 3 and 4 reported having a small amount of residual vision, however their performance when compared to the other BVI participants does not suggest this improved their performance in terms of worksheet accuracy or completion time.

6. Discussion

This study began with the question: *Can multisensory spatial inputs (high contrast visual, spatial language, and haptics) lead to the same level of learning for concepts that are conveyed through diagrams?* To investigate this question, we designed a multisensory learning system to evaluate its ability to deliver functionally equivalent spatial information (configuration and relationships) to communicate diagrammatic content. The solution addressed two primary considerations: (1) the multisensory system is based on a universal design approach providing spatial information in diagrams for all learners (including BVI learners) to participate in classroom activities (e.g., groupwork) with their peers – thus reducing barriers presented in the need to create separate, specialized materials for accommodations; and (2) a significant body of research has confirmed that information presented in complementary sensory modalities can enhance the acquisition and retention of information for all learners (i.e., benefit of multisensory information). We hypothesized that the multisensory interface would provide a highly similar (*functionally equivalent)* spatial information access experience for both sighted and BVI participants. The results corroborate this prediction suggesting that all participants received a similar level of spatial information through multisensory input channels that facilitated functionally equivalent communication and interpretation of the diagrams' content and meaning.

6.1. Worksheet accuracy

Our hypothesis that there would be no significant difference in *worksheet accuracy* performance between the MDS interface and control stimuli was supported by the null results, as there were no statistically significant differences observed in active-test worksheet accuracy based on participant groups. While there was a significantly lower overall accuracy performance for BVI participants as compared to sighted participants, this was only due to their significantly lower pre-test accuracy (prior knowledge), not active-test accuracy (learning gain). Comparing accuracy results across conditions (MDS and control) revealed a numerically small and statistically insignificant difference (4%) in active-worksheet accuracy between MDS and control conditions, with the mean information gain across conditions only differing by 1%. Therefore, although each group started with different levels of prior knowledge, their final scores (as measured by the active-test worksheet) were remarkably similar, supporting our *a priori* prediction in the ability of the MDS interface to provide the necessary spatial information using multisensory channels to lead to similar learning gains. A possible alternative hypothesis, where gains were only found for the control condition, would suggest that learning was possible but with differential performance between the experimental MDS condition and the standard haptic/visual modes. The absence of this finding, based on the highly similar performance on final active worksheets between conditions argues against this outcome and suggests the MDS was as effective in supporting knowledge gain. There was a significant effect of diagram type (atom diagram vs. atmosphere diagram), however, this was found across both participant groups, further supporting the similarity performance between the MDS and control conditions.

6.2. Worksheet time to completion

We hypothesized there would be no significant differences between participants for *worksheet completion* when using the MDS system and this assertion was also supported by the results. Time spent per question to complete the worksheet was not significantly different between MDS (45 s) and control (48 s) conditions or between groups. We hypothesized that increased geometric complexity of the individual diagrams (atmosphere v. atom) would increase worksheet completion times for all participants when using the MDS interface and that was indeed validated by the observed results, with the more complex diagram (atmosphere) taking significantly longer for both groups to interpret and answer questions.

The most notable outcomes of this study are the remarkably similar data in participant accuracy and completion time between modalities (MDS vs. Tactile and MDS vs. Visual Control), which provides compelling evidence of the effectiveness of this interface compared to the gold standard diagram rendering techniques and suggests that there was a high level of similarity in information gain. These results are especially promising given sighted users' lack of experience with vibro-audio information access. This outcome suggests that the multisensory channels of information can provide a functionally equivalent learning experience for students who may need different types of information to understand a complex diagrammatic register exchange (Dimmel and Herbst, 2015). The MDS used multisensory input to provide redundant content information about the diagrams' meaning and interpretation through different channels simultaneously. Thus, participants in both groups could interpret the diagram's spatial information (i.e., spatial configuration and relationships) using vibro-audio input with similar performance as they could using more familiar modalities (tactile or visual). The MDS was able to successfully communicate the type of information needed to complete the diagrammatic register interchange using a combination of information input.

Our findings suggest that given a well-developed multisensory system, such as the MDS prototype, most participants were able to interpret spatial information within a diagrammatic representation well enough to make sense of the graphics using the vibro-audio interface. While the findings support this is true for the simple stimuli used in this study, we acknowledge that further research is needed to investigate if this finding of equivalent performance would hold for more complex stimuli. In addition, the MDS system was effective for conveying non-visual information for working age adult learners with and without vision. This is an important finding as it represents a new universal design for learning approach for learners in a variety of STEM settings (e.g., college, vocational training, workplace professional development) to work cooperatively using the same reference materials and content platform, regardless of their visual status. The results also suggest this multisensory approach is viable as a multipurpose, affordable, mobile diagram display interface and as an accurate non-visual STEM graphical content learning tool. In the future, this type of MDS application could work in conjunction with diagram creation tutorials and an upload interface. As such, this multisensory approach addresses the long-standing challenge of providing consistent and timely access to accessible educational materials. With further development and testing, this type of system could have the important benefit of helping many learners with diverse learning needs who require additional multisensory supports from being left out of future STEM labor market opportunities due to a lack of adequate and accessible learning materials. These types of accessible STEM materials could help to improve low rates of STEM participation and career success by BVI students creating more accessible pathways for educational, employment, and lifestyle outcomes (Cryer et al., 2013; American Foundation of the Blind (AFB), 2017).

Our results provide further empirical support corroborating the growing body of evidence from multisensory learning demonstrating functionally equivalent performance. That is, when information is matched between inputs during learning, it provides a common level of access to key content, and the ensuing spatial image can be acted upon in an equivalent manner in the service of action and behavior, independent of the input source. Importantly, this study showed functional equivalence in two ways, similarity between learning inputs (i.e., the MDS vs. haptic and visual controls), and between participant groups (i.e., blind and sighted learners).

The finding of functional equivalence between our learning modalities is consistent with comparisons of these inputs (see Loomis et al., 2013 for review of this literature) but extends the theory to similar results in a new domain—interactions with STEM diagrams. The finding of equivalent performance between sighted and blind participants is also important as it supports the notion that when sufficient information is made accessible to these adult learners, they can perform at the same level as their sighted peers (Giudice, 2018). This outcome, as we observed here, speaks to the importance of providing accessible diagrams. However, as we discussed in the introduction, this access is not meant to support a specific population, providing information through multiple sensory modalities benefits all learners and is the cornerstone of good inclusive design. Indeed, we are all multisensory learners as this is how our brain works, taking in, learning, representing, and

TABLE 6 Summary of pre/post accuracy collapsed across diagram and mode.

Visual status	Pre-accuracy	Active-accuracy
BVI	13.7 [4.5–22.8]	96.1 [92.8–99.4]
Sighted	31.2 [20.1–42.4]	92.0 [88.5–95.4]

TABLE 7 Summary of pre/post accuracy collapsed across mode.

Visual status	Diagram	Gain	Pre-accuracy	Active- accuracy	Completion time
BVI	Atom	95.4 [90.1–100.8]	23.6 [7.4–39.7]	95.8 [91.0–100.6]	46.0 [31.8–60.2]
	Sky	96.2 [90.5–102.0]	3.7 [−1.9–9.4]	96.3 [90.6–101.9]	41.6 [31.4–51.7]
Sighted	Atom	79.2 [67.4–90.9]	62.5 [52.6–72.4]	93.1 [89.1–97.1]	65.7 [52.2–79.2]
	Sky	90.8 [84.9–96.7]	0 [0–0]	90.8 [84.9–96.8]	31.9 [28.1–35.7]

TABLE 8 Summary of pre/post accuracy collapsed across diagram.

Visual status	Condition	Gain	Pre-accuracy	Active-accuracy	Completion time
BVI	MDS	93.6 [87.5–99.6]	6.0 [−1.3–13.4]	94.0 [88.4–99.6]	38.2 [29.0–47.4]
	Control	98.1 [93.8–100.0]	21.3 [4.1–38.5]	98.1 [93.9–102.4]	49.4 [35.8–62.9]
Sighted	MDS	85.8 [77.8–93.8]	27.5 [11.6–43.4]	90.2 [84.8–95.6]	51.6 [37.5–65.6]
	Control	84.2 [73.0–95.3]	35.0 [18.1–51.9]	93.8 [89.1–98.4]	46.1 [34.8–57.3]

TABLE 9 Summary of BVI participant results by diagram and visual status.

ID #	Etiology of blindness	Residual vision	Mode	Diagram	Gain	Pre-accuracy	Active-accuracy	Completion time
1	Lebers congenital amaurosis	Light/ dark perception	MDS	Atom	100.0	12.5	100.0	52.0
			Control	Sky	100.0	0.0	100.0	32.2
2	Lebers congenital amaurosis	Light/dark perception	MDS	Sky	100.0	16.7	100.0	27.8
			Control	Atom	100.0	37.5	100.0	37.0
3	Pathological myopia	Light/dark perception in right eye, Fuzzy colors	MDS	Sky	83.3	0.0	83.3	36.7
			Control	Atom	100.0	50.0	100.0	59.0
4	Retinitis pigmentosa, atypical, with cone dystrophy	Light/dark perception, some functional peripheral	MDS	Atom	87.5	0.0	87.5	16.5
			Control	Sky	100	0.0	100.0	31.5
5	Retinitis Pigmentosa	Light/dark perception	MDS	Atom	83.3	25.0	87.5	51.8
			Control	Sky	100.0	16.7	100.0	47.2
6	Retinopathy of Prematurity	Light/dark perception	MDS	Sky	100.0	0.0	100.0	39.8
			Control	Atom	100.0	50.0	100.0	80.5
7	Glaucoma	Light/dark perception	MDS	Sky	100.0	0.0	100.0	33.7
			Control	Atom	100.0	37.5	100.0	31.6
8	Unknown	Light/dark perception	MDS	Atom	87.5	0.0	87.5	50.6
			Control	Sky	83.3	0.0	83.3	63.4
9	Congenital Cataracts, Glaucoma	Light/dark perception	MDS	Atom	100.0	0.0	100.0	34.8
			Control	Sky	100.0	0.0	100.0	61.8

acting upon information from multiple inputs in a seamless and integrative manner. The key role of multimedia and complementing sensory modalities has been shown to enhance the acquisition and retention of information in dozens of contexts and situations (Mayer, 2002). The current work builds on this literature. Our findings not only support the possibility of functional equivalence for STEM learning outcomes when an inclusive, universal-designed system is available, they also show that such multisensory interfaces benefit all learners and have the potential for many applications beyond traditional accessibility.

7. Limitations and future work

As this was a prototype designed for this study, there are limitations in the design of the current MDS system that could be improved with additional technical development. For instance, while the MDS application was designed for use with both vibration /haptic feedback, this component could be augmented and enhanced in future incarnations. New user interface (UI) elements being developed by our group and collaborators support new haptic profiles that would allow for a greater array of patterns, vibration styles, and haptic interactions with the MDS. Incorporating this development into future MDS design would allow for improved mapping of different diagram elements to haptic feedback. This would provide enhanced stimulus–response pairings that would likely both increase the type of information that could be presented through this modality and the overall efficiency of information encoding and learning strategy when using the MDS. In addition, work by our group and others on automating natural language descriptions could improve how key visual elements are conveyed through speech description when such annotations are created through an automatic vs. human-generated process. We also recognize the fact that the MDS may not be able to communicate other types of diagrams (e.g., charts, graphs, maps, etc) with the same level of effectiveness as the ones used in this study. We are in the process of running additional studies with new MDS features to explore the multisensory system's effectiveness with these additional types of visual representations. An additional consideration not addressed in this study is that while the MDS system was designed to support creation of accessible content, it still involves a significant amount of human intervention. Automating this process is a long-term goal of this project that would greatly streamline the creation of accessible content. In addition to the design limitations, our ability to differentiate among the groups was limited by the ceiling effect of our measurements. It is important to note that there was no reduction in performance across the comparison groups, but in future studies we plan to use more sensitive measures to investigate how variations in modality affect diagrammatic perception. Finally, future studies will need to evaluate the system with specific demographics (e.g., school/college aged people for classroom use, people in vocational settings for supporting work contexts, etc) to fully validate its use across learning settings.

8. Conclusion

The Multimodal Diagram System was designed with both sighted and BVI learners in mind. The goal of the MDS design was to create a STEM graphical content learning tool that could be used by all students to help facilitate communication and discussion between peoples with different visual abilities in a classroom. The results of this experiment provide clear support for the efficacy of our approach and of the MDS as a new, universally designed tool for providing inclusive STEM access for all.

Data availability statement

The datasets presented in this article are not readily available under IRB constraints. Requests to access the datasets should be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by University of Maine IRB, under application number: 2014-06-12. The participants provided their written informed consent to participate in this study.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This research was also funded by NSF grant IIS1822800 A Remote Multimodal Learning Environment to Increase Graphical Information Access for Blind and Visually Impaired Students (PI Giudice, Co-PIs Dimmel and Doore) and by NSF grant 2145517 CAREER: A Transformative Approach for Teaching and Learning Geometry by Representing and Interacting with Three-dimensional Figures (PI Dimmel).

Acknowledgments

We thank the University of Maine Center for Undergraduate Research (UMaine CUGR) and UMaine ASAP Labs for early funding and technical assistance in the development phase of the MDS system.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Abrahamson, D., Flood, V. J., Miele, J. A., and Siu, Y. T. (2019). Enactivism and ethnomethodological conversation analysis as tools for expanding universal Design for Learning: the case of visually impaired mathematics students. *ZDM* 51, 291–303. doi: 10.1007/s11858-018-0998-1

American Foundation of the Blind (AFB). (2017). *Interpreting bureau of labor statistics employment data*. Louisville, KY. Available at: http://www.afb.org/info/blindness-statistics/adults/interpreting-bls-employment-data/234

American Printing House for the Blind (APH) (1997). *Guidelines for design of tactile graphics*, Louisville, KY: APH.

American Printing House for the Blind (APH). (2019). *Annual Report 2019*. Louisville, KY: American Printing House for the Blind, Inc.. Available online at: http://www.aph.org/annual-reports.

Avraamides, M. N., Loomis, J. M., Klatzky, R. L., and Golledge, R. G. (2004). Functional equivalence of spatial representations derived from vision and language: evidence from allocentric judgments. *J. Exp. Psychol. Learn. Mem. Cogn.* 30:804. doi: 10.1037/0278-7393.30.4.804

Blaser, B., and Ladner, R. E. (2020). "Why is data on disability so hard to collect and understand?" in *2020 research on equity and sustained participation in engineering, computing, and technology (RESPECT)*, vol. *1* (Portland OR: IEEE), 1–8.

Brock, A., Truillet, P., Oriola, B., and Jouffrais, C. (2010). Usage of multimodal maps for blind people: why and how in ACM International Conference on Interactive Tabletops and Surfaces (November 2010) 247–248. Atlanta, GA: ACM. doi: 10.1145/1936652.1936699

Byrne, R., and Johnson-Laird, P. N. (1989). Spatial reasoning. *J. Mem. Lang.* 28, 564–575. doi: 10.1016/0749-596X(89)90013-2

Cairncross, S., and Mannion, M. (2001). Interactive multimedia and learning: realizing the benefits. *Innov. Educ. Teach. Int.* 38, 156–164. doi: 10.1080/14703290110035428

Chandler, D. (2007). *Semiotics: The basics*, New York, NY: Routledge.

Choi, S. H., and Walker, B. N. (2010). "Digitizer auditory graph: making graphs accessible to the visually impaired" in *CHI'10 extended abstracts on human factors in computing systems*, (Atlanta, GA: ACM), 3445–3450.

Cryer, H., Cryer, H., Home, S., Wilkins, S. M., Cliffe, E., and Rowe, S. (2013). *Teaching STEM subjects to blind and partially sighted students: Literature review and resources*, Birmingham, UK: RNIM Centre for Accessible Information.

Dimmel, J. K., and Herbst, P. G. (2015). The semiotic structure of geometry diagrams: how textbook diagrams convey meaning. *J. Res. Math. Educ.* 46, 147–195. doi: 10.5951/jresematheduc.46.2.0147

Doore, S. A., Dimmel, J. K., Xi, R., and Giudice, N. A. (2021). "Embedding expert knowledge: a case study on developing an accessible diagrammatic interface" in *Proceedings of the forty-third annual meeting of the north American chapter of the International Group for the Psychology of mathematics education*. eds. D. Olanoff, K. Johnson and S. Spitzer (Philadelphia, PA : Psychology of Mathematics Education), 1759–1763.

Ehrlich, K., and Johnson-Laird, P. N. (1982). Spatial descriptions and referential continuity. *J. Verbal Learn. Verbal Behav.* 21, 296–306. doi: 10.1016/S0022-5371(82)90626-0

Furió, D., González-Gancedo, S., Juan, M. C., Seguí, I., and Costa, M. (2013). The effects of the size and weight of a mobile device on an educational game. *Comput. Educ.* 64, 24–41. doi: 10.1016/j.compedu.2012.12.015

Giudice, N. A. (2018). "Navigating without vision: principles of blind spatial cognition," in *Handbook of behavioral and cognitive geography* (Northampton MA: Edward Elgar Publishing) doi: 10.4337/9781784717544.00024

Giudice, N. A., Betty, M. R., and Loomis, J. M. (2011). Functional equivalence of spatial images from touch and vision: evidence from spatial updating in blind and sighted individuals. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 621–634. doi: 10.1037/a0022331

Giudice, N. A., Klatzky, R. L., Bennett, C. R., and Loomis, J. M. (2013). Combining locations from working memory and long-term memory into a common spatial image. *Spat. Cogn. Comp.* 13, 103–128. doi: 10.1080/13875868.2012.678522

Giudice, N. A., Palani, H. P., Brenner, E., and Kramer, K. M. (2012). Learning non-visual graphical information using a touch-based vibro-audio interface in Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility, 103–110. Boulder, CO: ACM. doi: 10.1145/2384916.2384935

Goodwin, G. P., and Johnson-Laird, P. N. (2005). Reasoning about relations. *Psychol. Rev.* 112, 468–493. doi: 10.1037/0033-295X.112.2.468

Griffin-Shirley, N., Banda, D. R., Ajuwon, P. M., Cheon, J., Lee, J., Park, H. R., et al. (2017). A survey on the use of mobile applications for people who are visually impaired. *J. Vis. Impair. Blindn.* 111, 307–323. doi: 10.1177/0145482X1711100402

Hake, R. R. (1998). Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses. *Am. J. Phys.* 66, 64–74. doi: 10.1119/1.18809

Hasty, L., Milbury, J., Miller, I., O'Day, A., Acquinas, P., and Spence, D. (2011) *Guidelines and standards for tactile graphics: Technical report, braille Authority of North America*, 2011. Available online at: http://www.brailleauthority.org/tg/

Heward, W. L. (2003). *Exceptional children: An introduction to special education*, New Jersey: Pearson Education.

Honey, M., Alberts, B., Bass, H., Castillo, C., Lee, O., and Strutches, M. M. (2020). *STEM education for the future: a visioning repor*t. National Science Foundation Technical Report, 36.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*, Cambridge, MA: Harvard University Press, 6.

Johnson-Laird, P. N. (2006). "Mental models, sentential reasoning, and illusory inferences" in *Advances in psychology*, vol. *138* (North-Holland), 27–51. doi: 10.1016/S0166-4115(06)80026-9

Johnson-Laird, P. N. (2010). Mental models and human reasoning. *Proc. Natl. Acad. Sci.* 107, 18243–18250. doi: 10.1073/pnas.1012933107

Johnson-Laird, P. N., and Byrne, R. M. (1991). *Deduction,* Hove, East Sussex, United Kingdom: Lawrence Erlbaum Associates, Inc..

Kent, M., Ellis, K., Latter, N., and Peaty, G. (2018). The case for captioned lectures in Australian higher education. *TechTrends* 62, 158–165. doi: 10.1007/s11528-017-0225-x

Khine, M. (2013). *Critical analysis of science textbooks. Evaluating instructional effectiveness*. Netherlands: Springer. doi: 10.1007/978-94-007-4168-3

Klatzky, R. L., Lippa, Y., Loomis, J. M., and Golledge, R. G. (2003). Encoding, learning, and spatial updating of multiple object locations specified by 3-D sound, spatial language, and vision. *Exp. Brain Res.* 149, 48–61. doi: 10.1007/s00221-002-1334-z

Knauff, M., and Ragni, M. (2011). Cross-cultural preferences in spatial reasoning. *J. Cogn. Cult.* 11, 1–21. doi: 10.1163/156853711X568662

Kress, G., and van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (*2nd ed.*). London, England: Routledge. doi: 10.4324/9780203619728

Loomis, J. M., Klatzky, R. L., and Giudice, N. A. (2013). "Representing 3D space in working memory: spatial images from vision, hearing, touch, and language" in *Multisensory imagery: Theory & Applications*. eds. S. Lacey and R. Lawson (New York: Springer), 131–156. doi: 10.1007/978-1-4614-5879-1_8

Lundgard, A., and Satyanarayan, A. (2021). Accessible visualization via natural language descriptions: a four-level model of semantic content. *IEEE Trans. Vis. Comput. Graph.* 28, 1073–1083. doi: 10.1109/TVCG.2021.3114770

Mayer, R. E. (2002). "Multimedia learning" in *Psychology of learning and motivation*, vol. *41* (New York, NY: Academic Press), 85–139. doi: 10.1016/S0079-7421(02)80005-6

McDonnall, M. C., and Sui, Z. (2019). Employment and unemployment rates of people who are blind or visually impaired: estimates from multiple sources. *J. Vis. Impair. Blindn.* 113, 481–492. doi: 10.1177/0145482X19887620

Moon, N. W., Todd, R. L., Morton, D. L., and Ivey, E. (2012). *Accommodating students with disabilities in science, technology, engineering, and mathematics (STEM).* Atlanta, GA: Center for Assistive Technology and Environmental Access, Georgia Institute of Technology.

Moreno, R., and Mayer, R. E. (2002). Verbal redundancy in multimedia learning: when reading helps listening. *J. Educ. Psychol.* 94, 156–163. doi: 10.1037/0022-0663.94.1.156

Moreno, R., and Mayer, R. (2007). Interactive multimodal learning environments. *Educ. Psychol. Rev.* 19, 309–326. doi: 10.1007/s10648-007-9047-2

National Center for Education Statistics (NCES) (2022). "Students with disabilities" in *Condition of education* (U.S. Department of Education, Institute of Education Sciences. Available online at: https://nces.ed.gov/programs/coe/indicator/cgg)

National Federation of the Blind (NFB). *Blindness statistics*. (2009). Available at: https://nfb.org/blindness-statistics

Palani, H. P., Fink, P. D. S., and Giudice, N. A. (2020). Design guidelines for schematizing and rendering Haptically perceivable graphical elements on touchscreen devices. *Int. J. Hum. Comput. Interact.* 36, 1393–1414. doi: 10.1080/10447318.2020.1752464

Palani, H.P., and Giudice, N.A. (2016). *Usability parameters for touchscreen-based haptic perception. In 2016 IEEE Haptics Symposium (HAPTICS).* (Philadelphia, PA, USA: IEEE).

Ragni, M., and Knauff, M. (2013). A theory and a computational model of spatial reasoning with preferred mental models. *Psychol. Rev.* 120, 561–588. doi: 10.1037/a0032460

Schles, R. A. (2021). Population data for students with visual impairments in the United States. *J. Vis. Impair. Blindn.* 115, 177–189. doi: 10.1177/0145482X211016124

Schneiderman, B., Plaisant, C., Cohen, M., and Jacobs, S. (2018). *Designing the user Interface: Strategies for effective human-computer interaction*, *6th ed.*; Addison Wesley: Boston, MA, United States.

Sharif, A., Chintalapati, S. S., Wobbrock, J. O., and Reinecke, K. (2021). Understanding screen-reader users' experiences with online data visualizations in In Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21), Association for Computing Machinery, New York, NY, USA. 1–16.

Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J. K., Guild, N., et al. (2009). Why peer discussion improves student performance on in-class concept questions. *Science* 323, 122–124. doi: 10.1126/science.1165919

Tipton, K. (2021). *Video captions as a universal design technique: impact on student success*. California State University: Northridge.

Trickett, S. B., and Trafton, J. G. (2006). "Toward a comprehensive model of graph comprehension: making the case for spatial cognition" in *International conference on theory and application of diagrams* (Springer, Berlin, Heidelberg: Springer), 286–300.

U.S. Census Bureau (2020). *2019 American community survey 1-year estimates: Employment status by disability status and type, Table B18120*. Available online at: https://data.census.gov/cedsci/tableq=B18120&tid=ACSDT1Y2019.B18120

Ungar, S., Blades, M., and Spencer, C. (1996). "The construction of cognitive maps by children with visual impairments" in *The construction of cognitive maps* (Dordrecht: Springer), 247–273. doi: 10.1007/978-0-585-33485-1_11

Vickers, P., Faith, J., and Rossiter, N. (2012). Understanding visualization: a formal approach using category theory and semiotics. *IEEE Trans. Vis. Comput. Graph.* 19, 1048–1061. doi: 10.1109/TVCG.2012.294

Voltz, D. L., Brazil, N., and Ford, A. (2016). What matters most in inclusive education. *Interv. Sch. Clin.* 37, 23–30. doi: 10.1177/105345120103700105

W3C (2019). Web Content Accessibility Guidelines (WCAG) http://www.w3.org/TR/WCAG/W3C is the publisher). *WAI web accessibility tutorials: Complex images*

WebAim.org (2019). *Survey on screen reader Usage #9*. Available online at: https://webaim.org/projects/screenreadersurvey9/

Yatani, K., Banovic, N., and Truong, K. (2012). SpaceSense: representing geographical information to visually impaired people using spatial tactile feedback in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Austin, TX: ACM. 415–424.

Yelland, N. J. (2018). A pedagogy of multi-literacies: young children and multimodal learning with tablets. *Br. J. Educ. Technol.* 49, 847–858. doi: 10.1111/bjet.12635

Zhang, M. R., Zhong, M., and Wobbrock, J. O. (2022). "Ga11y: an automated GIF annotation system for visually impaired users," in *Proceedings of the ACM conference on human factors in computing systems (CHI'22)* (New Orleans, New York, United States: ACM), 1–16.

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 17 April 2023
DOI 10.3389/feduc.2023.1023499

Check for updates

OPEN ACCESS

EDITED BY
Wang-Kin Chiu,
The Hong Kong Polytechnic University,
Hong Kong SAR, China

REVIEWED BY
Mawardi Mawardi,
Padang State University, Indonesia
Ririn Ambarini,
Universitas PGRI Semarang, Indonesia

*CORRESPONDENCE
Aliashim Albani
a.albani@umt.edu.my

SPECIALTY SECTION
This article was submitted to STEM Education, a section of the journal Frontiers in Education

RECEIVED 19 August 2022
ACCEPTED 14 March 2023
PUBLISHED 17 April 2023

CITATION
Gopabala Krishnan L, Albani A, Segar Y, Ridzuan AR, Ismail MK, Musa MA, Gunaseelan K, Subramaniam TS and Ibrahim MZ (2023) The student response on the use of renewable energy graphical interface simulator in learning environment. *Front. Educ.* 8:1023499. doi: 10.3389/feduc.2023.1023499

The student response on the use of renewable energy graphical interface simulator in learning environment

Lavaniya Gopabala Krishnan[1], Aliashim Albani[2]*, Yogesswary Segar[1], Abdul Rahim Ridzuan[3], Mohd Khairi Ismail[4], Mohd Azlan Musa[2], Khemaalatha Gunaseelan[5], Tamil Selvan Subramaniam[5] and Mohd Zamri Ibrahim[2]

[1]Faculty of Business, Economics and Social Development, Universiti Malaysia Terengganu, Kuala Nerus, Malaysia, [2]Renewable Energy & Power Research Interest Group (REPRIG), Faculty of Ocean Engineering Technology Informatics, Universiti Malaysia Terengganu, Kuala Nerus, Malaysia, [3]Faculty of Business and Management, Universiti Teknologi MARA, Melaka, Malaysia, [4]Faculty of Business and Management, Universiti Teknologi MARA, Terengganu, Malaysia, [5]Faculty of Technical and Vocational Education, Universiti Tun Hussein Onn Malaysia, Johor, Malaysia

Introduction: The purpose of this study is to evaluate students' responses regarding the usefulness of a graphical user interface (GUI) tool in the context of their respective learning environments. The Energy Computator GUI (EC-GUI) helps to simplify the STEM student's learning processes. The EC-GUI serves as a simulator that can assist in computing formulas, designing graphs, acting as a unit converter, and automatically deriving parameters.

Methodology: Furthermore, a survey, which included closed and open questions, was carried out on a selection of students majoring in STEM subjects at Universiti Malaysia Terengganu (UMT) who were enrolled in the Renewable Energy course. A total of 54 respondents participated in the survey and 90.8% of them expressed satisfaction with the EC-GUI provided. The research involved using two distinct kinds of analysis: a parametric analysis, the paired sample *t*-test, and a non-parametric analysis, the Wilcoxon signed-rank test.

Results: The study findings indicated that the majority of the respondents felt that the difficulty level of the subjects did not change after using the EC-GUI. However, it helped to simplify the learning process for students in STEM fields. The *p*-value of the appropriate teaching aid tool was less than 0.05, indicating that the results were significant both before and after using the EC-GUI.

Conclusion: The study suggests that a similar GUI tool could be implemented in Malaysia's teaching and learning processes as it is easy to build and use.

KEYWORDS

STEM, teaching, learning, and assessment, graphical interface simulator, education, Malaysia

1. Introduction

STEM is an acronym for four closely related fields of study: science, technology, engineering, and mathematics. STEM is a challenging subject and STEM courses are disliked by many students (John and Estonanto, 2017). STEM education has received much interest over the last decade (Honey et al., 2014; Alam et al., 2021). STEM education combines ideas that are often taught independently in different courses and focus on the ability to apply knowledge to actual issues. As a result, the world has need of additional STEM-literate individuals, experts, and leaders, and its significance cannot be overemphasized (Alam et al., 2021). Moreover, STEM also refers to all technologies that use mathematics and science (Albani and Ibrahim, 2019). Author explains that non-effective teaching and learning processes encountered by students in the classroom might be one of the factors for the drop in interest in STEM courses.

The integration of Information, Communication, and Technology (ICT) can support educators in meeting the necessity of replacing conventional teaching approaches with technology-based teaching and learning resources and infrastructures (Ghavifekr and Rosdy, 2015). Technology is an essential topic in many sectors, including education, in the twenty-first century, as technology has become the preferred method of information transfer in most countries. Technology-based teaching and learning may bring about numerous improvements in the classroom, but good planning and policy development are required (Ghavifekr and Rosdy, 2015).

Assessment is a key component of the teaching and learning process because it gathers, interprets, and analyses student progress data. The effectiveness of assessment techniques in the classroom determines the quality of learning. The difficulties in the education system require the development and execution of ideas of teaching, learning, and assessment (TLA) methods that consider both rational thoughts of knowledge and practical implementations while also decreasing the time needed to perform the TLA cycle (Albani and Ibrahim, 2019). An extremely imaginative and unique strategy is essential to ensure students comprehend the knowledge shared during lectures (Hussin et al., 2017). An approach involving immersive online and offline materials can increase students' inspiration toward understanding advanced courses, allowing them to acquire high grades (Hussin et al., 2017).

A graphical user interface (GUI) is a graphical representation of one or more windows containing commands or features that enable users to execute activities effectively. The simulator defines as a type of user interface in which people interact with digital equipment using graphical icons and visual indications (Nass et al., 2021). This visual system design is not intended to replace traditional programming but rather to convert pictorial presentations into structural lists, after which researchers generate a GUI that represents as a communication interface throughout which users can interact with the computer graphically (Kholil and Wahyudin, 2018). A GUI can also be defined as an application demonstrating menu options, icons, navigation, and other tools that replace command prompts or shell commands (Kholil and Wahyudin, 2018). Moreover, the study focuses on the development of the Energy Computator GUI., or EC-GUI, a teaching aid tool to address a renewable energy technology and energy management challenge. The study seeks to present the notion of a check and balance approach (CBA) in teaching, learning, and evaluation processes. The designed EC-GUI tool is forecast to expand the teaching quality, learning, and evaluation of comprehending both the computation algorithms and theories.

The major contribution of the study is to enhance the importance of using better teaching aid tools for difficult subjects, such as in STEM studies. STEM subjects are unpopular among high school graduates, and most of the students deny entry to STEMs due to complex subjects and also a less proper ways to identify the terms and equation. This study aims to help boost students' motivation to work smarter by using the GUI interface to accumulate such terms and equations. The EC-GUI is designed to be intuitive and simple to use, making it ideal for use in educational settings. The application, along with other ways of solving difficulties, will help to improve high school students' ability to study STEM subjects.

The research focuses on students' feedback toward the teaching aid tool and the efficiency of the learning process using the EC-GUI. The study objectives are: (1) to develop a TLA aid tool for the course of Renewable Energy Technology and Energy Management, and (2) to analyze the impact of the developed teaching aid tool in the learning process. The study was done with a closed- and open-ended survey evaluation. It was carried out with a limited sample of enrolled students from the Renewable Energy Technology and Energy Management course. The selected candidates were STEM students from Universiti Malaysia Terengganu (UMT). The analysis of the study focused on parametric and non-parametric methods. The parametric method used the paired sample ***t***-test, and the non-parametric used the Wilcoxon signed-rank test. The total number of respondents were 54 students who were enrolled in the Renewable Energy course. The outcome of the study parameters were significant. The difficulty of the course was kept the same or constant, but the teaching aid tool helped to simplify the lectures. The students were pleased with the teaching aid tool, and the ***p***-value of the test was less than 0.05 both before and after. The parametric and non-parametric test were validated, reliable, and significant regarding the sample collected. The results proved that the hypothesis of the study was achieved in the creation of the EC-GUI teaching aid tool for STEM students and the subject of Renewable Energy Technology and Energy Management.

2. Benchmark of simulation tool

2.1. The graphical interface development

The MATLAB-based learning aid tool was developed for the purpose of enhancing students' comprehension in order to improve their ability to solve a specific energy-related design problem. A graphical user interface, also known as a GUI, is an interactive display consisting of one or more windows and various controls and components that allow users to complete tasks interactively. The developed interface, which can be seen presented in **Figure 1**, is made up of a number of different components that are utilized in the process of carrying out a simulation. These components can take many forms, such as menus, buttons, tables, and axes. The EC-GUI tool was developed with the intention of being able to solve design problems in the following fields (as shown in **Figure 2**): (i) Renewable Energy Technology; (ii) Energy Management; and (iii) Energy and Climate Change. The user of the tool, who may be a student or a teacher, operates the tool and manipulates the control element in order to carry out an act of simulation. The linear and non-linear equation solvers provided in the MATLAB optimization toolbox function are used for any computation that needs an iterative process to solve. This function is adopted by the MATLAB optimization toolbox. The result is produced at the

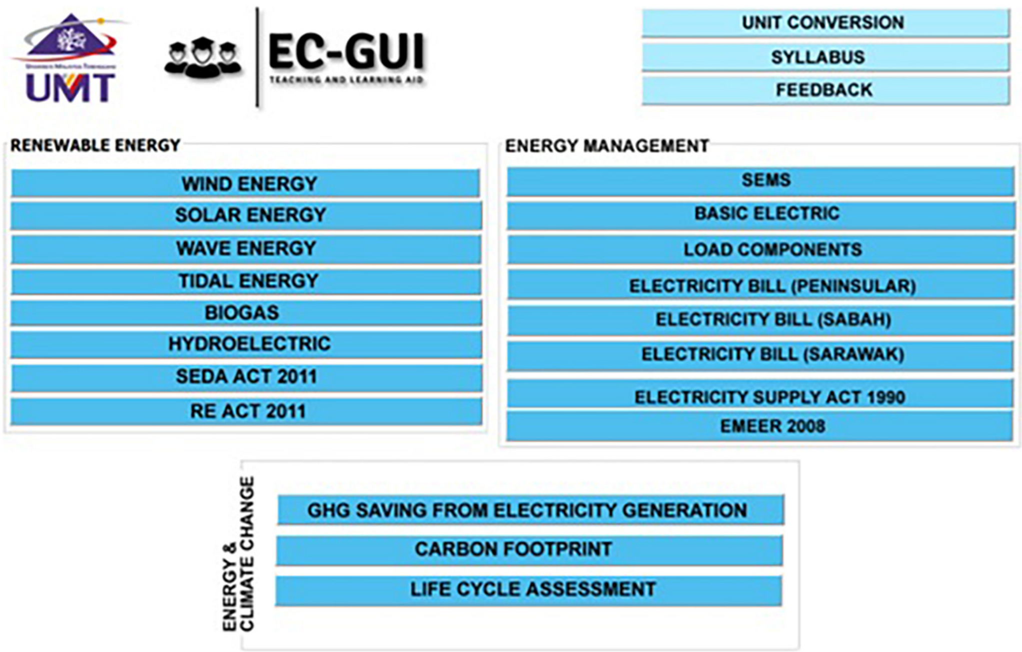

FIGURE 1
The main screen of the graphical user interface (GUI) panel.

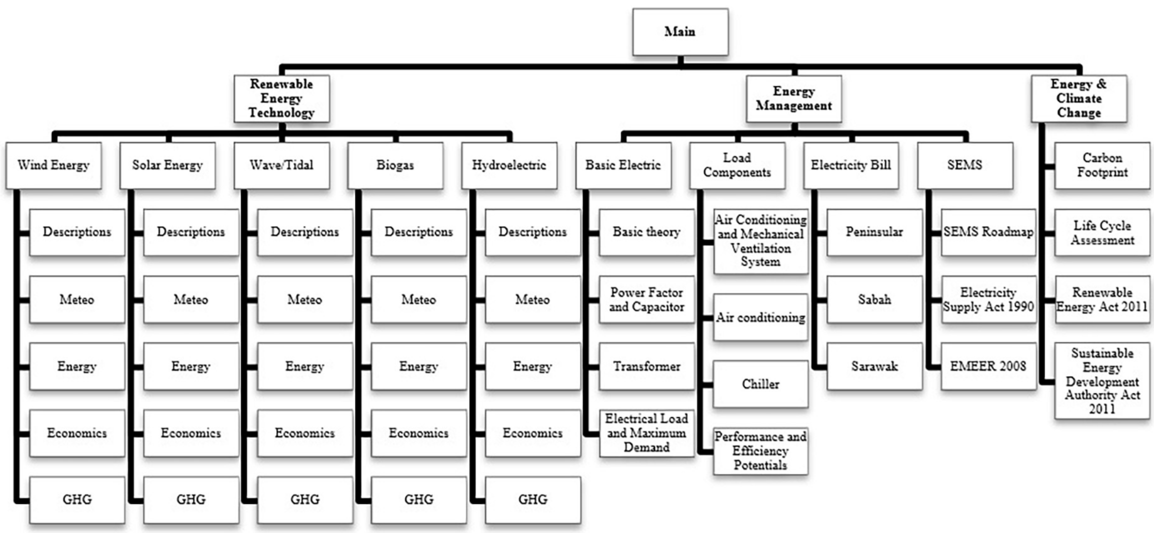

FIGURE 2
The GUI tool map.

conclusion of the simulation process, and it is presented both numerically and graphically.

3. Application in teaching learning and assessment

The designed EC-GUI interface is an interactive application that allows users to be fully involved in a simulation and discovery learning. Overall, a check and balance approach (CBA) was designed to shorten and simplify the TLA procedure. According to the Oxford English Dictionary (OED), the word "check" is defined as the process of examining something to see if it is safe, correct, and acceptable (Stevenson, 2010); the word "balance" is defined as maintaining things in equal, correct, and the right proportions, also according to the OED (Stevenson, 2010). As a result, the terms check and balance combined were added to the TLA approach, in which both the educator and student participate to comprehend the underlying ideas and properly solve any graphical interface design issues (Albani and Ibrahim, 2019). **Figure 3** below explains the flow of the CBA within the TLA approach.

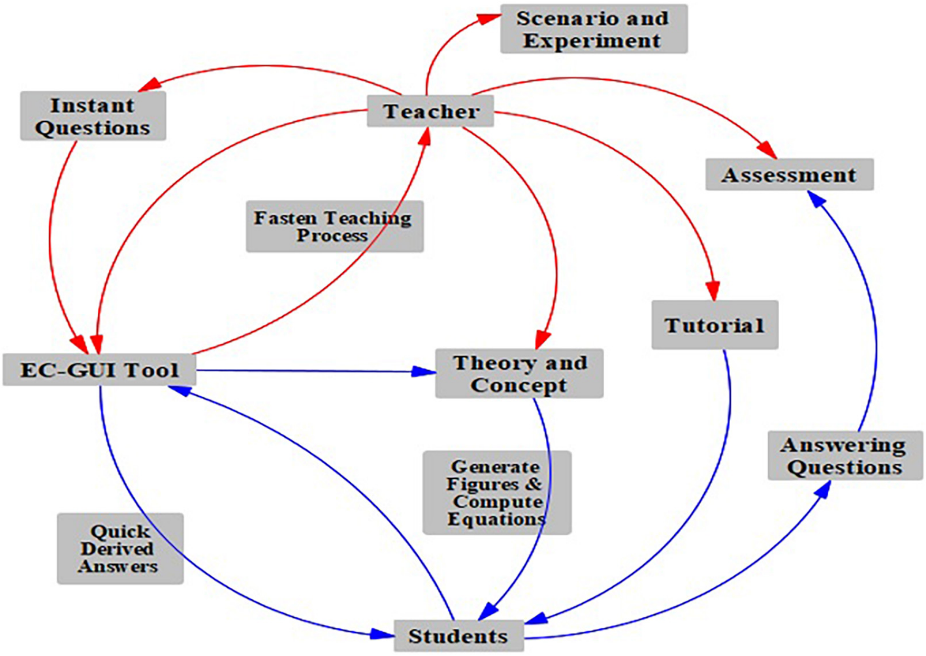

FIGURE 3
The check and balance approach.

There are three key components in the CBA approach: (1) the graphical interface tool (EC-GUI), (2) the teaching, learning, and assessment (TLA) process, and (3) the users (teacher/instructor or students). For instructors, the EC-GUI might be utilized as a teaching tool for Renewable Energy and Energy Management course concepts and theories, as well as for tutorial class teaching. Furthermore, the teacher can use the EC-GUI as a calculator to speed up the evaluation of students' assignments and exams. Similarly, for students, the EC-GUI can be used as a learning tool for course ideas and theories, as well as for calculations in tutorial classes. Students can even use it as a calculator to ensure the accuracy of their answers when completing assignments or self-learning.

The EC-GUI application tool has advantages with a CBA in the TLA process: by using iterative calculations and sensitivity analysis, the tool will improve the comprehension of the Renewable Energy Technology and Energy Management subject's notions or theories. The EC-GUI application also lets students monitor performance by comparing iterated outcomes in various ways. Therefore, students can improve their understanding of Energy Management and Renewable Energy Technology more efficiently by using this technology and also assess the impact of various Energy subject unit operations. In essence, when the tool is applied in the TLA process, there is a double-feedback process, enabling the user to examine the multiple design capabilities of the EC-GUI unit process.

After evaluating the global results of the unit process performance, the users may draw numerous assumptions, after which the users can adjust the parameters and repeat the simulation. This repeated process improves students' understanding of the EC-GUI design and operation. Furthermore, the designed technology enables students to double-check their work while self-learning and finishing projects. Students will benefit from the GUI tool's self-discovery capability, including in understanding and addressing their unique problems during learning process. Also, the proposed tool assists in reducing the time required to finish the TLA procedure, particularly during short semesters.

TABLE 1 Five point likert scale.

Scale value	Level of likert scale
1	Strongly disagree
2	Disagree
3	Moderate
4	Agree
5	Strongly agree

4. Research methods

4.1. Close-ended survey analysis

The close-ended survey was conducted among the STEM students enrolled in the Renewable Energy Technology and Energy Management course. The total number of respondents was 54 students. The questionnaire was distributed to the 54 students and consisted of multiple sections. The questionnaire was split into two elements: preliminary action and the final action of the research, which meant there will be two different questionnaires with the same questions and sections. The questionnaires were conducted before using the EC-GUI tool,

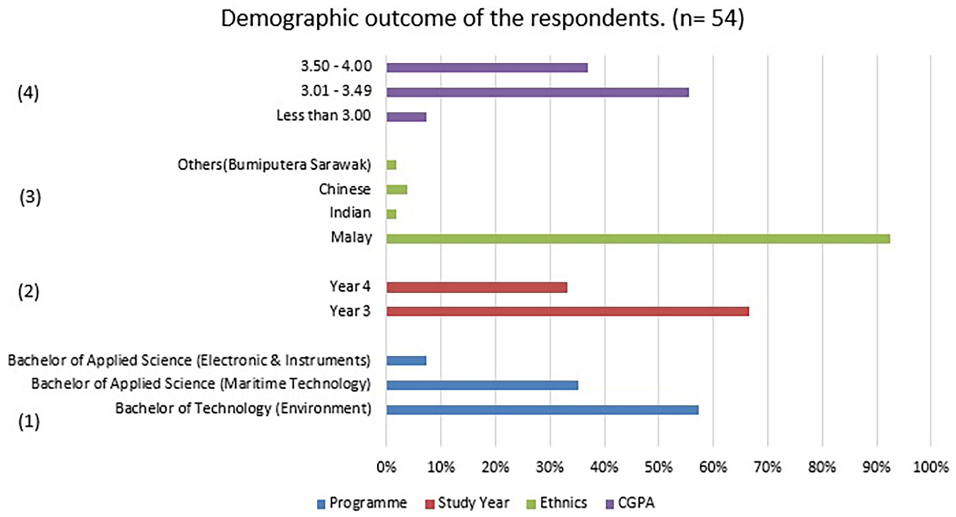

FIGURE 4
The demographic result of the user (respondents) of teaching aid tool.

and then the following questionnaires were conducted after utilizing the EC-GUI tool. This was to help identify the efficiency of the tool. The survey consisted of four parts: (1) Common section: Demographics, (2) Part A: Level of difficulties for energy subjects, (3) Part B: Relationship of student interests with props, and (4) Part C: Characteristics of appropriate props. The questionnaire consisted of 30 questions. The demographic part of the survey was analyzed using descriptive analysis. Part 2 and part 3 were analyzed using both parametric and non-parametric analysis, which helped to compare the pre- and post-study results of the survey. Likert scales were used in all of the questions to assist analysis of every parameter. Further explanation regarding the analyses used are listed below with precise descriptions.

4.1.1. Likert scale

The Likert scale is a psychometric measure widely used during survey research. The Likert scale has been utilized in all the sections of the questionnaire. **Table 1** explains the levels of the scales used.

4.2. Parametric analysis: Paired sample *t*-test

Parametric analysis is a statistical analysis in which a test establishes certain requirements regarding the form of distribution of parameters or populations, such as whether the data are interval-scaled and normally distributed (Yosani, 2006). The term "parametric analysis" refers to statistical procedures that are predicated on the premise that the distribution of population data is normal (Nawangsari, 2017).

This study focused on the paired ***t***-test in regard to the interval scales of the data. The paired sample ***t***-test was used to examine whether the mean difference of the data should be adjusted before performing parametric tests (Khosravi et al., 2018). As a result, the authors conducted a descriptive study before doing the ***t***-test. When there was only one group of people and we needed to gather data from them about two distinct times, we used the paired sample ***t***-test, and if the sig(2-tailed) value was less than 0.5, there was a significant difference between the two scores (Khosravi et al., 2018).

The paired sample *t*-test, also known as the dependent sample *t*-test, is a statistical process for determining if the mean difference between two groups of data is zero. Each subject or object is measured twice in a paired sample *t*-test, resulting in pairs of observations. Studies and repeated-measures techniques are two common uses of the paired sample *t*-test. The below equation represents the paired *t*-test (Boyd, 2020).

$$t = \frac{\overline{Xd} - \mu_o}{S_D/\sqrt{n}}$$

4.3. Non-parametric analysis: Wilcoxon signed-rank test

Non-parametric analysis is used for data that are devoid of distribution *(Teguh, 2014)*; non-parametric analysis is statistical analysis that does not specify normally distributed data conditions *(Yosani, 2006)*. Non-parametric statistics use statistical approaches that do not assume that population parameter distribution is normal *(Nawangsari, 2017)*.

Like the related *t*-test, the Wilcoxon test can be used as a non-specific test of the empirical hypothesis, 'the empirical measure differs in the two samples', or as a particular test of an empirical shift hypothesis (Kornbrot, 1990). The Wilcoxon is "appropriate" in circumstances with non-normally distributed interval measures because it has the highest asymptotic relative efficiency for exploiting the information provided in the sample ranks (Kornbrot, 1990).

The non-parametric form of the paired samples *t*-test is the Wilcoxon signed-rank test. When the distribution of the differences

TABLE 2 Non-parametric test using Wilcoxon signed-rank test (Part A).

	Questions	Z	Asymp. sig. (2-tailed)
Part A: levels of difficulties for energy subjects	1. I am interested in studying the subject of energy.	−1.013	0.311
	2. The subject of energy is a difficult subject to excel in (Grade A).	−0.561	0.575
	3. I have problems solving the calculation parts of energy subjects.	−0.012	0.991
	4. I have problems understanding the theories of the subject of energy.	−1.331	0.183
	5. I often postpone energy subject assignments or projects until the last minute due to lack of understanding.	−0.564	0.573
	6. I feel energy subjects are very complex as well as difficult to learn.	−1.877	0.061
	7. I feel the topics containing equation problems in energy subjects are the most difficult to study.	−2.251	0.024**
	8. I feel it is difficult to understand conventional teaching techniques in the classroom.	−0.221	0.825
	9. I am not interested in conventional or passive teaching techniques in the classroom.	−0.172	0.864

**Significant at five percent.

between the two samples cannot be assumed to be normal, it is used to determine whether there is a significant difference between two population means. The test's purpose is to assess if two or more sets of pairings differ from one another in a statistically meaningful way. Equation 2 below shows the formula of Wilcoxon signed-rank test (MacFarland and Yates, 2016).

$$W = \sum_{i=1}^{N} [sgn\,(x2, i - x1, i) \times R_i$$

5. Results and discussion

5.1. Demographic information of the respondents

Section "5.1. Demographic information of the respondents" emphasizes the demographic profiles of the respondents. The respondents of the study were students at UMT. **Figure 4** summarizes the overall findings as well as the demographic information of the respondents. The demographics profiles contain the respondents' program of study, year of study, ethnicity, and cumulative grade point average (CGPA). The results indicated that most of the respondents (57.4%) were from the Bachelor of Technology (Environment) program and in their third year of study (66.7%). Regarding ethnicity, as can be seen in **Figure 4**, the majority of respondents identified as Malay, with all other ethnicities reported, such as Indian, only accounting for 1.9% combined. The lowest number of students whom performed in the survey is others category which is Bumiputera Sarawak. Moreover, **Figure 4** shows the CGPA of the students specifically in ranges. The highest number of students were categorized in the 3.01 to 3.49 CGPA range and the lowest were in the range of less than 3.00 CGPA. The demographic sections contributed the details of the respondents clearly in a few questions. For further details, see **Figure 4**.

5.2. Analysis using the non-parametric Wilcoxon signed-rank test

In the case that the normality assumption was not satisfied, we used the Wilcoxon signed-rank test to perform the comparison on the medians. In SPSS, we used Analyze, then Non-parametric tests, then Two related samples: **Tables 2–4** show the SPSS output. A $\boldsymbol{p}$-value < 0.05 was considered significant for all statistical tests conducted. The Z reported in **Tables 2–4** is the test statistic for the non-parametric Wilcoxon signed-rank tests. It is a statistic that is used in hypothesis testing for the non-parametric Wilcoxon signed-rank test and is the sum of the signed ranks. This test was conducted on: the levels of difficulties for energy subjects (Part A), the relationship of students' interest with props (Part B), and the characteristics of appropriate props (Part C), which were all taken from the survey data.

A student survey for the levels of difficulties for energy subjects (Part A) was implemented pre and post use of the teaching aids tool (EC-GUI). The Wilcoxon signed-rank test results indicated a statistically significant difference between the perceived pre and post levels of difficulty for energy subjects. Since the p-value (0.024) was less than 0.05, we can reject the null hypothesis (see **Table 2**). We have sufficient evidence to conclude that "the topics containing equation problems in energy subjects are the most difficult to study" is statistically significant, meaning the difficulties of the subjects became easier after using the EC-GUI.

Further, the difference between the pre and post scores for the relationship of students interests with props (see **Table 3**) was also statistically significant in "The teaching aid tools used during the teaching were able to maintain my interest in the subject of energy" with a p-value of 0.030, "My interest in the subject of energy increased when the lecturer used the teaching aid tools in his teaching" with a p-value of 0.046, "I am interested in doing energy subject calculation exercises provided by lecturers" with a p-value of 0.038, and "I am interested in doing energy subject exercises that are available in reference books other than those given by the lecturers after using the teaching aid tools" with a p-value of 0.021. These results explain that the teaching aid tool (EC-GUI) assisted students in increasing their interest in the learning process and cultivated

TABLE 3 Non-parametric test using Wilcoxon signed-rank test (Part B).

	Questions	Z	Asymp. sig. (2-tailed)
Part B: relationship of students' interest with props.	10. I am very interested in learning the subject of energy when lecturers use Teaching Aid Tools while teaching.	−1.250	0.211
	11. I don't like studying energy subjects without Teaching Aid Tools.	−0.612	0.541
	12. I get bored of learning energy subjects if the lecturers do not use Teaching Aid Tools during teaching.	−1.417	0.157
	13. I enjoy learning the subject of energy with Teaching Aid Tools.	−1.623	0.105
	14. The use of Teaching Aid Tools by the lecturers has become a motivation for me to follow the subject of energy.	−0.847	0.397
	15. I cannot concentrate even though the lecturer used Teaching Aid Tools while teaching.	−0.757	0.449
	16. I enjoy learning the subject of energy if the lecturers use Teaching Aid Tools.	−1.403	0.161
	17. The Teaching Aid Tools that are used during the teaching maintained my interest in the subject of energy.	−2.169	0.030**
	18. My interest in the subject of energy increased when the lecturer used Teaching Aid Tools in his teaching.	−1.992	0.046**
	19. I am interested in doing energy-subject calculation exercises provided by lecturers.	−2.075	0.038**
	20. The use of Teaching Aid Tools by lecturers did not directly interest me in the subject of energy.	−0.457	0.648
	21. Learning the subject of energy is not interesting with the use of Teaching Aid Tools.	−0.348	0.728
	22. The use of Teaching Aid Tools by lecturers has sparked my interest to be more diligent in studying the subject of energy.	−1.098	0.272
	23. Learning the subjects of energy becomes more interesting with the availability of Teaching Aid Tools.	−1.427	0.154
	24. I am interested in doing energy subject exercises that are available in reference books other than those provided by the lecturers after using the Teaching Aid Tools.	−2.313	0.021**
	25. I managed to solve energy subject problems when the lecturers used Teaching Aid Tools while teaching.	−1.015	0.310

**Significant at five percent.

TABLE 4 Non-parametric test using Wilcoxon signed-rank rest (Part C).

	Questions	Z	Asymp. sig. (2-tailed)
Part C: characteristics of appropriate props.	26. The Teaching Aid Tools used by lecturers in accordance with the content of learning delivered.	−2.645	0.008**
	27. The Teaching Aid Tools used by the lecturers can be clearly seen.	−2.912	0.004**
	28. The Teaching Aid Tools used by the lecturers caught my attention.	−2.747	0.006**
	29. The Teaching Aid Tools used by lecturers are organized and easy to use.	−1.935	0.053**
	30. I am pleased if the following Teaching Aid Tools are used by the lecturers while teaching:		
	a. Printed material	−1.063	0.288
	b. Non-printed material	−2.100	0.036**
	c. 3D material	−0.145	0.885
	d. Video	−2.214	0.027**
	e. Audio	−0.525	0.599
	f. Software or mobile apps	1.22	0.222

**Significant at five percent.

their spirits to learn about the subject of energy. Regarding the outcome of the Part B section, the students were highly interested in doing calculations using smoother tools like the EC-GUI interface tool.

The Wilcoxon signed-rank test result indicated a statistically significant difference between the pre and post scores for appropriate teaching aid tools characteristics, as seen in **Table 4**. **Table 4** shows the multiple types of teaching aid tools used in the classes, whereas the EC-GUI interface is the most highly relevant tool for STEM students especially for energy subjects. We found statistically significant improvements in the characteristics of appropriate EC-GUI tools in terms of the responses to question

TABLE 5 Overall outcome of the paired *t*-test for the pre and post surveys.

	Paired difference							
				95% confidence interval of the difference				
	Mean	Std. deviation	Std. error mean	Lower	Upper	t	df	Sig. (2-tailed)
Part A total of pre and post findings	0.68519	3.28992	0.44770	0.21279	1.58316	1.530	53	0.132
Part B total of pre and post findings	2.37037	7.75156	1.05485	4.48614	25460	2.247	53	0.029**
Part C total of pre and post findings	5.51852	4.46682	0.60786	6.73773	4.29931	9.079	53	0.000**

**Significant at five percent.

(26) Teaching aid tools were used by lecturers in accordance with the content of learning delivered, (27) Teaching aid tools could be clearly seen, (28) Teaching aid tools caught my attention, (29) Teaching aid tools were organized and easy to use, and (30) Teaching aid tools were used by the lecturers while teaching especially non-print materials and videos. The levels of significance are clearly shown in Table 5. The teaching aid tools such as non-print materials and videos were highly attractive to STEM students; non-print materials had a ***p***-value of 0.036 and videos had a ***p***-value of 0.027. These two teaching tools were highly accepted by students as elements of teaching. Other tools, such as audio, printed materials, and 3D materials, were not as highly regarded by the students. Non-printed materials and videos can be used later on after the class for further discussions. This result revealed that the teaching aid tools (GUI) made difficult subjects easier, especially those associated with many calculations, formulas, theories, and concepts.

5.3. Pre and post survey analysis: Parametric analysis—Paired sample *t*-test

In this section, we discuss the paired sample *t*-test between the average means of constructs obtained from the pre and post survey data regarding: the levels of difficulties for energy subjects (Part A), the relationship of students' interests with props (Part B), and characteristics of appropriate props (Part C). For this purpose, we computed the paired sample *t*-test between the pre and post means of the teaching aid tool (EC-GUI) as follows.

A paired sample *t*-test was performed between the levels of difficulty for the subject of energy (Part A) pre and post using the teaching aids tool (EC-GUI). Table 5 shows the results of the paired sample *t*-test. The mean difference was computed as 0.68519 with a standard deviation of 3.28992, whereas the t-statistics was 1.530, the degrees of freedom (df) was 53, and the sig (2-tailed) *p*-value was 0.132. These results meant there was no significant difference in the levels of difficulty for the subject of energy before and after using the teaching aid tool. Further, a paired sample *t*-test was also performed between the relationship of students' interests with the teaching aid Tool (Part B) before and after using the teaching aid tool (EC-GUI). The result of the paired sample *t*-test showed the mean difference was computed as 2.37037 with a standard deviation of 7.75156, whereas the t-statistics was 2.247, the degrees of freedom (df) was 53, and the sig(2-tailed) *p*-value was 0.029. These results meant there was a significant relationship between students' interests in the teaching aid tool before and after using it and that the teaching aid tool (EC-GUI) made it easier and more efficient to understand the calculation approaches and subject concepts [t(53) = 2.247, $p \leq 0.05$]. Additionally, a paired sample *t*-test of the characteristics of appropriate teaching aid tools (Part C) was carried out. The paired *t*-test results, as presented in Table 5, indicated that the difference between the pre and post means of using teaching aid tools was 5.51852, with a standard deviation of 4.46682; furthermore, the t- statistics was 9.079, the degrees of freedom (df) was 53, and the sig(2-tailed) *p*-value was 0.000, meaning there was a significant difference in the characteristics of appropriate teaching aid tools and the teaching aid tools were more affordable as compared to the before [t(53) = 9.0749, $p \leq 0.05$].

The results indicated that the teaching aid tool is efficient enough and STEM students preferred using the EC-GUI teaching aid tool. The tool was sufficient to help the students become more interested in energy subjects despite STEM subjects being categorized as difficult subjects by the majority of the students. Whereas the tools of multiple features help the students to study in more comfortable ways. Regarding using the calculator for mathematical calculations, STEM students, especially students studying energy subjects, need advanced technology and software such as EC-GUI to be more efficient. Additionally, the response to the initial query also agreed with author (4) in that the teachers of the specified subject should value the use of diverse tactics that stray from the traditional teaching of the subject. Here, the teaching aid tools need to be smarter and more helpful for the students, such as the EC-GUI tool. The pre survey results of the study explained the difficulties of students to adapt to STEM subjects, especially energy subjects. The necessity to tackle engineering issues or mathematics in the syllabus frequently contributes to students' perceptions of the subject's difficulty (4). Here, the post survey showed how satisfied the students were after using the EC-GUI.

6. Conclusion

The overall research shows that EC-GUI teaching aid tool is highly efficient among STEM students. However, 63% of students strongly agreed that the Teaching Aid Tools that the lecturers had used while teaching in the current semester. The EC-GUI did not reduce the difficulties of the subject but it did help to enhance the ways to learn more efficiently. It helped the students to work on energy subjects in a much simpler manner. The main objective of the teaching aid tool is to provide more flexible ways of learning and implementing equations and the study proved that STEM students were able to use it to study in a wiser manner. Hence, in the future more students or STEM candidates should use teaching aid tool interfaces as an easier way to accumulate the outcomes of functions and graphic images. The parametric and non-parametric results indicated that the students preferred to use tools such as the EC-GUI. The study aimed to estimate the efficiency of the EC-GUI teaching aid tool among STEM students associated together with a reduction in the perceived complicatedness of energy subjects. The research was conducted among UMT STEM students enrolled in energy subjects. This teaching aid tool will help to improve students' capabilities to derive formulas and compute equations and graphs easily. Altogether, the study's findings indicate that the designed tool is a possible tool that may help in the teaching, learning, and evaluation processes of the challenging subject of energy taught in a higher-learning institution. To sum up, using the EC-GUI tool in the teaching, learning, and assessment (TLA) process helps the understanding of energy subjects and contributes a few advantages:

(i) The EC-GUI tool can be used in TLA processes with easy functions for the tutors.
(ii) The designed interface can be quickly installed on any computer without the need to acquire a database or costly software.
(iii) The users can use the tool to derive formulates, compute formulas, and show figures and graphs.
(iv) The users are allowed to change the settings of the EC-GUI tool according to self-preferences to produce the desired results.

Limitations can be identified in multiple scope in a study. Here, the graphical user interface (GUI) was found by the authors to bring a smoothness to the learning processes of STEM subjects. The EC-GUI application has multiple uses for the students and this is also a limitation in the study where this application was only used by students from UMT. Several research gaps exist in this field, given the need for more research on STEM education in Malaysia. The study gap defines that STEM regarding research are rare. Research in future can use the same EC-GUI interface to evaluate the difficulties of STEM subjects in a wider range of survey analyses by allocating more funds for its usage. The research can also be conducted with more variety of student's perceptions, by adding other characteristics into the EC-GUI interface. This could justify the usage of the teaching aid tool as an efficient necessity for STEM students. In the future, researchers can use a wider sampling selection by having a greater number of respondents. In this study, the number of students for the pre and post study survey were the same.

Data availability statement

The original contributions presented in this study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by the Universiti Malaysia Terengganu Research Ethics Committee (No: UMT/JKEPM/2022/100). The patients/participants provided their written informed consent to participate in this study.

Author contributions

AA conceived and designed the investigation. LGK and YS drafted the report and had principal responsibility for the interpretation of data. MKI, ARR, MAM, KG, TSS, and MZI critically reviewed and revised the study and provided important intellectual and educational content. All authors take responsibility for the final approval of the version to be published and are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved, read, and agreed to the published version of the manuscript.

Funding

This study was funded by the Universiti Malaysia Terengganu through the Scholarship of Teaching of Learning (SoTL), grant number (VOT. 55199/1).

Acknowledgments

Permission to conduct the investigation was obtained following review by the Universiti Malaysia Terengganu Research Ethics Committee (No: UMT/JKEPM/2022/100).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Alam, M., Sajid, S., Kok, J., Rahman, M., and Amin, A. (2021). Factors that influence high school female Students' intentions to pursue science, technology, engineering and mathematics (STEM) education in Malaysia. *Pertanika J. Soc. Sci. Humanit.* 29, 839–867. doi: 10.47836/pjssh.29.2.06

Albani, A., and Ibrahim, M. (2019). Development of graphical interface simulator of advanced wastewater treatment design process for teaching, learning, and assessment. *Designs* 3, 1–11. doi: 10.3390/designs3020027

Boyd, R. (2020). "PART 3 advanced topics" in *Tolerance Analysis of Electronic Circuits Using MATLAB*, CRC Press, Taylor & Francis Group, 127–128. doi: 10.1201/9780203735619

Ghavifekr, S., and Rosdy, W. (2015). Teaching and learning with technology: effectiveness of ICT integration in schools. *Int. J. Res. Educ. Sci.* 1, 175–191. doi: 10.21890/ijres.23596

Honey, M., Pearson, G., Schweingruber, H., and National Academy of Engineering and National Research Council (2014). *STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research*, Vol. 10. Washington, DC: National Academies Press, 18612.

Hussin, H., Abidin, H., Husaini, Y., Muhamad, M., and Ali, M. (2017). "Students' perception on teaching and learning activities and their understanding in power engineering course," in *Proc 2017 IEEE Int Conf Teaching, Assess Learn Eng TALE 2017*, Hong Kong, 333–338. doi: 10.1109/TALE.2017.8252358

John, A., and Estonanto, J. (2017). Acceptability and difficulty of the STEM track implementation in senior high school. *Asia Pacific J. Multidiscip Res.* 5, 43–50.

Kholil, M., and Wahyudin, D. (2018). Graphical user interface (GUI) of GPS data (Dumy Data). *IOP Conf. Ser. Mater Sci. Eng.* 384:012070. doi: 10.1088/1757-899X/384/1/012070

Khosravi, M., Ali, N., Karbasi, M., Brohi, I., Shaikh, I., and Shah, A. (2018). Comparison between NFC/RFID and bar code systems for Halal tags identification: paired sample T-test evaluation. *Int. J. Adv. Comput. Sci. Appl.* 9, 221–229. doi: 10.14569/IJACSA.2018.090435

Kornbrot, D. (1990). The rank difference test: a new and meaningful alternative to the Wilcoxon signed ranks test for ordinal data. *Br. J. Math Stat. Psychol.* 43, 241–264. doi: 10.1111/j.2044-8317.1990.tb00939.x

MacFarland, T., and Yates, J. (2016). *Introduction to Nonparametric Statistics for the Biological Sciences Using R*. Berlin: Springer. doi: 10.1007/978-3-319-30634-6

Nass, M., Alégroth, E., and Feldt, R. (2021). Why many challenges with GUI test automation (will) remain. *Inf. Softw. Technol.* 138:106625. doi: 10.1016/j.infsof.2021.106625

Nawangsari, T. (2017). "Test of difference between paired sample by using Mcnemar test", in *International Conference on Mathematics: Education, Theory, and Application (ICMETA)* 2017, 117–125. Available online at: https://jurnal.uns.ac.id/math/article/viewFile/12036/10558 (accessed December 31, 2022).

Stevenson, A. (2010). *Oxford Dictionary of English*. Oxford: Oxford University Press.

Teguh, M. (2014). *Metode Kuantitatif untuk Analisis Ekonomi dan Bisnis*. Depok City: PT. RajaGrafindo Persada, 234.

Yosani, C. (2006). *Teknik Analisis Kuantitatif. Makal Tek Anal II*. 1–7. Available online at: http://staffnew.uny.ac.id/upload/132232818/pendidikan/Analisis+Kuantitatif.pdf (accessed December 29, 2022).

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 20 April 2023
DOI 10.3389/feduc.2023.1150190

Check for updates

OPEN ACCESS

EDITED BY
Wang-Kin Chiu,
The Hong Kong Polytechnic University,
Hong Kong SAR, China

REVIEWED BY
Nicoleta Duta,
University of Bucharest, Romania
Hideyuki Kanematsu,
National Institute of Technology, Suzuka College, Japan

*CORRESPONDENCE
Daniela Galatro
daniela.galatro@utoronto.ca

SPECIALTY SECTION
This article was submitted to STEM Education, a section of the journal Frontiers in Education

RECEIVED 23 January 2023
ACCEPTED 30 March 2023
PUBLISHED 20 April 2023

CITATION
Chakraborty S, Gonzalez-Triana Y, Mendoza J and Galatro D (2023) Insights on mapping Industry 4.0 and Education 4.0. *Front. Educ.* 8:1150190. doi: 10.3389/feduc.2023.1150190

Insights on mapping Industry 4.0 and Education 4.0

Sourojeet Chakraborty[1], Yris Gonzalez-Triana[2], Jorge Mendoza[3] and Daniela Galatro[1]*

[1]Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON, Canada, [2]Facultad de Ciencias Naturales y Matemáticas, Escuela Superior Politécnica del Litoral (ESPOL), Guayaquil, Ecuador, [3]Facultad de Ingeniería en Ciencias de la Tierra, Escuela Superior Politécnica del Litoral (ESPOL), Guayaquil, Ecuador

Introduction: The fourth industrial revolution, or Industry 4.0 (I.D. 4.0), has radically empowered professionals to revamp skills and technologies, to match ever-evolving industry demands. Education 4.0 (E.D. 4.0) is an integral education framework, strategically designed to align with I.D. 4.0 needs. The present work presents high-level insights on mapping I.D. 4.0 to E.D. 4.0, by successfully analyzing the four key existing components of E.D. 4.0, namely, learning methods, competencies, infrastructure and information and communication technologies (ICT).

Methods: Research questions are formulated along themes aiming to standardize the E.D. 4.0 framework and identify effectiveness and implementation challenges. These posed questions are addressed by performing an exhaustive bibliometric analysis on the associated literature, by clustering relevant publications by field, year, and geography. We employed the search engines Scopus, Science Direct, and IEEE in a period between January and June of 2022.

Results: Network maps evidence the implementation of E.D. 4.0 elements with no formal and universally adopted framework to map with I.D. 4.0. There is an increasing interest and support from researchers and education institutions in preparing a skilled workforce for I.D. 4.0. Trends of E.D> 4.0-related published articles reveal more implementation efforts in developed countries compared to developing countries.

Discussion: Our results demonstrate a lack of any currently existent, standardized, and universally accepted framework for mapping I.D. 4.0 to E.D. 4.0, despite trends showing a sharp rise towards incorporating E.D. 4.0 initiatives recently into university curricula. Our analysis procedure can serve as a protocol to define E.D. 4.0 in a more specific context, in an ever-changing global workspace. While unbalanced implementation attempts on how extensively E.D. 4.0 components have been defined and adopted (including discrepancies in implementation policies among countries, and across disciplines), further rigorous assessments are needed to critically assess the necessary requirements and effectiveness, for standardization and implementation a global mapping framework.

KEYWORDS

Education 4.0, Industry 4.0, mapping, systematic literature reviews, bibliometric analysis, pedagogy, engineering education, higher education

1. Introduction

Professionals are expected to continue developing skills, technologies, and knowledge, adaptable toward assimilating fast-paced, swiftly changing innovations, driven by the fourth industrial revolution or Industry 4.0 (hereafter, I.D. 4.0), which impact products, services, and production systems (Bauer et al., 2015; Hirschi, 2018; Kipper et al., 2021). Strategic attempts toward identifying these key competencies, as required by future professionals, have been reported in recent years; and has resulted in the identification of "competency clusters." From an extensive literature review, Hecklau et al. (2016) grouped these competencies into four clusters: technical, methodological, social, and personal. A more recent survey conducted across industries and industry representatives (Mian et al., 2020), have classified clusters by disciplines (engineering, business, design, and transversal). Yet, another survey-based study (Hernandez-de-Menendez et al., 2020) performed across several industries, interpreted these necessary trends to be predominantly technology-driven, and recognized particularly paradigm advances in the field of Artificial Intelligence (AI), owing to the increasing globalized availability of big data, which in turn, propels industry growth. ChatGPT, an AI based chatbot, which was launched in November 2022, has quickly attracted global hype for its promise of enhancing both supervised and reinforcement-based learning techniques across both industry and academia. Despite having some obvious current limitations (such as monitoring and calibration of factual inaccuracies), the parent firm (OpenAI) has been valued at US $29 billion, as of 2023. An overwhelming majority of recent technologies incorporated into the present global industrial market (expectedly) values Machine Learning (ML) as the most utilized skill (87%), followed by user/entity big data analytics (84%), the Internet of Things (IoT) (82%), and cloud computing (76%) (AnTosz, 2018). A major finding from this work is that skills gaps in the local labor market accounts for 59% of the barriers that exist toward the adopting of new technologies. Therefore, it is the need of the hour to swiftly revamp global education to address I.D. 4.0 driven innovation(s), and expand worker competencies/skills, to match industry requirements (Neaga, 2019; Uhlemann et al., 2019; Kipper et al., 2021).

The necessary integral education framework that aligns with I.D. 4.0 requirements is Education 4.0 (hereafter, E.D. 4.0). Several components and defined characteristics are captured in the E.D. 4.0 framework; these have been identified rather recently by some prior researchers (Fisk, 2017; Himmetoglu et al., 2020; Kipper et al., 2021; Miranda et al., 2021). It is imminent that such a standardized pathway needs to be identified and incorporated to approach such problems, as a current lack of standardization negatively impacts the reliability, consistency, and reproducibility of any findings. An unreliable process is difficult to properly analyze, and any associated bias at any step of the analysis procedure ultimately delays a tangible mapping of E.D. 4.0 techniques with I.D. 4.0 requirements, further augmenting an already existent, and progressively increasing, skilled labor shortage.

Even today, there appears to be no universally agreed upon definition or terminology of E.D. 4.0 (Das et al., 2020). Thus, to address the need for having a standardized protocol that can identify, define and shape an effective E.D. 4.0 pedagogical framework, and identify key elements, we performed an exhaustive literature review on the currently existent I.D. 4.0 competencies and compared several approaches, ultimately choosing to merge the most up to date factors identified by Kipper et al. (2021) and Miranda et al. (2021). Once merged, we hypothesized the global need for obtaining a standard E.D. 4.0 definition, as the currently incomplete, scattered framework for I.D. 4.0 will very likely lead to scattered solutions when mapped with I.D. 4.0. A major highlight of our work is to identify and track the most recent E.D. 4.0 competencies, which allows the analysis procedure to be as generic as possible, to ensure maximal success and universality for future mapping efforts. From a comparative study of these competencies, four research questions (RQs) are formulated, which are centered along themes of the extent of successful implementation of the E.D. 4.0 framework, its effectiveness (both perceived and actual), and its associated challenges. Therefore, the objectives that characterize our proposed framework are: (a) identifying the extent to which E.D. 4.0 is presently adopted as a formal educational framework for mapping with I.D. 4.0 requirements, (b) identifying the extent to which E.D. 4.0 components have been already adopted from the formal E.D. 4.0 framework present in the relevant literature; and (c) analyzing implementation trends and challenges of the E.D. 4.0 framework and its associated core components globally, over the last decade. Previous research in this sector has always been limited by the very fact that I.D. 4.0 requirements change over time as industries evolve globally (Wallner-Drewitz and Wagner, 2016): therefore, we recognize that E.D. 4.0 competencies must also change with time to map I.D. 4.0 needs. Thus, to succinctly capture these competences and associated insights, we performed a rigorous bibliometric analysis using VOSviewer, on several research databases, which provided straightforward, simple, objectively unbiased, reliable indicators of the impact, importance, and emerging future trends of this research field. In addition to shaping the proposed E.D. 4.0 framework, our findings are also valuable toward identifying curriculum changes and current pedagogical lacunae, such as modifications in existing course syllabi, the design of new, relevant, strategic programs and courses, etc. Such initiatives are extremely likely to guide the commensurate acquisition and targeted training of precious human capital to drive industry desired changed, as one is easily able to track the evolution and effectiveness of E.D. 4.0 implementation. This paper is structured into the following sections: section "2. Literature review" presents the currently existent E.D. 4.0 framework, and its associated indicators and competencies via a detailed literature review, section "3. Materials and methods" presents the mapping strategy and methodology employed to complete the integral E.D. 4.0 framework, formulates the relevant RQs arising out of this integrated framework, describes the relevant protocols employed toward our bibliometric analyses, and describes the platform that was employed to create these network maps, arising from the bibliometric data. Section "4. Results" presents our results, and section "5. Discussion" thoroughly interprets the results obtained, providing detailed discussion, context, and trends on the visualization and exploration of network maps and graphs, which provides answers to our previously postulated RQs. From the conclusions arising from bibliometric analyses, we propose a more holistic, universally applicable, and representative (and, updated) definition of E.D. 4.0, that is most applicable to the current world. Finally, in section "6. Conclusion," we present our conclusions, and identify scope for future research work toward mapping E.D. 4.0 to I.D. 4.0 needs in the future.

2. Literature review

Efforts to comprehensively map E.D. 4.0 to I.D. 4.0 needs, have predominantly focused on approaches that typically integrate competency and capability function domains (Wilke and Magenheim, 2017; Das et al., 2020; Jerman et al., 2020; Maisiri and van Dyk, 2020; Agrawal et al., 2021; Silva et al., 2021). Survey-based approaches (Grzybowska and Łupicka, 2017) have identified major E.D. 4.0 components to be decision making, entrepreneurial thinking, efficiency orientation, problem solving, conflict resolution, and analytical skills (in decreasing rank order). More recently, a more "integrated alignment model" was proposed (Lin and Low, 2021) to capture a more synergistic alignment between educational activities, and the I.D. 4.0 demands, as applicable to the Singapore Smart Industry Readiness Index (SIRI). In their approach, road mapping and architecture planning for progressive phases for the integrated alignment model, was linked to four key pillars: connectivity, automation, operation, and intelligence. This exercise was jointly performed alongside an industrial partner, thereby allowing for the identification of E.D. 4.0 tasks per phase. While this approach is definitely synergistic, we hypothesize that in fact, a more clustered approach toward identifying E.D. 4.0 competencies (rather than approaching from a more task-based format), allows for more efficient mapping between these two areas, while also capturing their shared dependencies. From our hypothesis, it follows that the definition and clustering of E.D. 4.0 competencies and their subsequent mapping with I.D. 4.0 requirements, must be approached using a more systematic format; one that strategically aims to list/map them, and then, standardize these requirements cohesively and comprehensively. It is evident that I.D. 4.0 must serve as the starting point for the mapping process, as skills, requirements, and demands of I.D. 4.0 are much more robustly defined by employers and industries (and are also, constantly evolving). In this context, we have already successfully incorporated a specific I.D. 4.0 demand, of understanding "Standards, Codes, and Recommended Practices," in the chemical engineering sector, at our university (Galatro et al., 2022). Owing to the ever-changing nature of I.D. 4.0 needs, it is not surprising that some prior researchers (Beke et al., 2020) have attempted to identify these competencies by conducting detailed interviews with industries, and surveys with students, to list some identifiable I.D. 4.0 requirements and expectation. Interestingly, some shared requirements also emerged from the expectations of students and industries; these are: complex problem solving, coordinating with others, people management, critical thinking, negotiation, quality control, service orientation, judgment and decision making, active listening and creativity (ranked in 2020, in decreasing order of perceived importance). Eight of these parameters were also ranked important in 2015 (with the exception of quality control and active listening). As the authors themselves state, this work (while commendable) suffers from some limitations, such as the lack of industry representation (interviews limited to the car industry), geographic limitation (student interviews were conducted only at one university in Hungary) and a drastic lack of representation among disciplines (survey limited to mechanical engineering). Additionally, there was no attempt to systematically categorize any identified competencies, which reinforces the need for analyzing this problem through a more structurally rigorous yet sound lens.

Hernandez-de-Menendez et al. (2020) have identified key competencies for I.D. 4.0, by analyzing various models that assess the maturity and readiness of companies to shift to I.D. 4.0 frameworks. Through a global-centric assessment, identified competencies were grouped into three broad categories: methodological, social, and personal. But it must be noted that these competencies were identified only for three disciplines (engineering, business, and design); therefore, no insight about desired competencies in other disciplines such as science, humanities, etc. were obtained. The work concludes by recognizing that there remains no universal consensus on required I.D. 4.0 competencies, and there may in fact be other more systemic economic/social barriers at play which prevent its successful incorporation (firms may choose not to shift toward I.D. 4.0, fearing higher economic expenditure; and employees may prefer not to adapt to newer, ever changing industry demands).

More recent characteristics of I.D. 4.0 have been identified more recently by Mian et al. (2020), as customization, real-time monitoring, productivity, flexibility, logistics, product design/prototyping, resource allocation, responsiveness, sustainability, process reliability, and predictive maintenance (and there may be many others that could exist, some of these characteristics may pertain to more specific industries). An attempt to map I.D. 4.0 has recently been performed by Maisiri and van Dyk (2020), who developed a Competence Maturity Model (CMM), with three domains for competency, capability functions, and maturity level. Each domain has two dimension which encompasses both knowledge and skills (technical vs. soft) requirements. While the capability domain has ten dimensions (all related to industrial engineering), and the maturity level has five dimensions (in line with the industrial revolution); all of which were identified by a systematic mapping review of 283 published papers (out of which only 25 papers were included for the purpose of data abstraction). While clearly conceptualized, well-supported, and efficiently implemented, the competencies are exclusively limited to industrial engineering, thereby preventing its global application across other fields/disciplines.

A more thorough (and recent) literature review has been credited to González-Pérez and Ramírez-Montoya (2022) who propose eight key (RQs); and we also follow a similar approach when formulating our RQs in section "3.1. Hypotheses and research questions" to identify components that use the E.D. 4.0 framework, from 113 reports (out of these, 56 are finally analyzed). The data predicts a major shift occurring in pedagogical practices, with case studies and targeted teaching/learning strategies to gain prominence over the 21st century. Learning methods and competencies are found as the most addressed components of E.D. 4.0, while a scarcity of frameworks are identified, which aim to address strategies to strengthen pedagogical innovation, especially at the school level. These limitations, in turn, become the merits of this work, namely: (a) the selection of a robust framework to assess core components in educational initiatives/projects, and (b) the identification of trends in the identified competencies for various crucial players such as researchers, trainers, and decision makers.

The initiatives and projects in this study were evaluated using the core components framework, defined by Miranda et al. (2021). The work successfully develops a comprehensive E.D. 4.0 framework, based on four critical components: competencies, learning methods, infrastructure, and information and communication technologies (ICTs). This framework arises from a compilation of research elements, structures, and concepts across several infographic sources that align with E.D. 4.0 concepts (as summarized in **Tables 1**, **2**). However, the case studies reported by Miranda et al. (2021) are localized to Mexico City and surveyed an English/Spanish-speaking student body; and therefore, suffer from geographical and linguistic restrictions.

A summary of some important prior works, with their employed methodologies, and limitations toward mapping these competencies effectively, are presented in **Table 1**. We have focused predominantly on experimental/literature review-based works, which have attempted to characterize and/or map these two factors. While the works listed represent a conscious academic effort toward mapping, **Table 1** also justifies the need to perform a more exhaustive bibliometric analysis, which can overcome the inherent/identified limitations in these prior works. For our bibliometric analysis, the Inclusion and Exclusion criteria is kept globally applicable, to ensure that our results are applicable in the broadest possible context, thereby paving the way for pursuing more rigorous, detailed studies in the future.

An extremely detailed approach toward identifying E.D. 4.0 competencies related to the qualification of professionals for I.D. 4.0 has been recently performed by Kipper et al. (2021) by surveying the literature with the SciMAT scientific mapping software on the Scopus, Web of Science, and Science Direct databases. The mapping resulted in the generation of a conceptual map, highlighting the major competencies (leadership, strategic vision of knowledge, self-organization, offering and receiving feedback, pro-activity, creativity, problem-solving, initiative, interdisciplinary teamwork, collaborative teamwork, innovation, communication, adaptability, flexibility, and self-management) and knowledge (information and communication technology, algorithms, automation, software development and security, data analysis, general systems theory, and sustainable development theory) required for the successful transformation of firms toward I.D. 4.0 targets. While this exhaustive search omits "learning methods" and "infrastructure" as core components; it nevertheless identifies some essential elements that may be clustered in the competencies and ICTs components, proposed by Miranda et al. (2021). The work of Kipper et al. is also much more universally applicable, as the bibliometric analysis is not limited to a specific discipline, and therefore, the conclusions hold much more universally. **Table 2** presents a summary of E.D. 4.0 competencies, as identified by some prior works.

It appears that a very impactful strategy toward the identification of any existent relationships and/or mapping efforts toward linking E.D. 4.0 to I.D. 4.0 is bibliometric analysis (Janik and Ryszko, 2018), and this is the approach we resort to in this work. With the advent of technology, in an increasingly digitized world, it is strategic to resort to computer-based analysis techniques. To ensure that the articles matching the first bibliometric criterion are indeed accurate and relevant, it is customary to refine the results further, to obtain the most accurate results, by implementing an inclusion and exclusion criteria. While this is definitely a robust procedure, which leads to successive refinement of data; it often results in a rather small final set of article database to base analysis/comparison/conclusions on. A summary of past bibliometric analyses performed on the E.D. 4.0/I.D. 4.0 literature, and the effective number of articles finally analyzed, are presented in **Table 3**.

The universalization of E.D. 4.0 remains yet another unaddressed challenge, in addition to a lack of standardization in defining and applying an E.D. 4.0 framework, which could then successfully map out I.D. 4.0 requirements. As of 2018, the level of I.D. 4.0 implementation for developing nations was captured at the corporate level, with strategies adopted by separate countries (Bogoviz et al., 2019); and at the national level for developed nations, with state-based strategies of development. Furthermore, major financial barriers exist on the path toward I.D. formation and their consequential implementation, as well as a gap in terms of the readiness of various socio-economic platforms toward the formation of I.D. 4.0 (Costan et al., 2021). Expectedly, the results in ultimately forming I.D. 4.0 targets currently reveal a 5-year gap between developed and developing countries. Such imbalance in implementation strategies must also be considered by researchers, when future attempts to map E.D. 4.0 to I.D. 4.0 are implemented, as specific systemic barriers exist, that hinder the universalization of the E.D. 4.0 framework. Currently, extremely limited works have been conducted to identify these E.D. 4.0 implementation barriers for developing economies, using the number of peer-reviewed publications as an effort indicator, both during the COVID-19 pandemic, or before. In fact, the recent COVID-19 pandemic has categorically highlighted the stark inequality and resource discrepancies between developing and developed nations, at both economic and social strata (Perry et al., 2021; Wakamo, 2022); this has also directly affected education systems. Some other barriers toward E.D. 4.0 implementation for developing economies are the lack of appropriate ICT infrastructure and widespread access; these might significantly deepen inequality. We also recognize the need for rigorous works to assess the global impact of these outcomes, as the United Nations (UN) Sustainable Development Goal (SDG) 4, which ensures quality education for all, might never be realized unless we take imminent swift action in this field.

3. Materials and methods

3.1. Hypotheses and research questions

Our literature review reveals no existent unanimous consensus toward accurately defining a current E.D. 4.0 framework. This lack of a standardized framework results in the existence of several scattered approaches/solutions toward mapping I.D. 4.0, which creates more imbalance toward tackling these lacunae, from a strategy-oriented perspective. Following a deep dive into the literature, we propose to merge the frameworks of Kipper et al. (2021) and Miranda et al. (2021), since both studies are: (a) fairly exhaustive, (b) based on systematic literature reviews, and (c) extremely recent, thereby ensuring up-to-date completeness of required competencies, since the fields of E.D. 4.0 and I.D. 4.0 continually evolve. We refer to this combined framework as the "reference framework" for our analysis. The reference framework

10.3389/feduc.2023.1150190

TABLE 1 A summary of key prior attempts by researchers, to map E.D. 4.0 to I.D. 4.0, and limitations.

References	Methodology	Comments/insights/limitations
Bauer et al., 2015	Quantitative literature review to track methods to realize I.D. 4.0.	Analysis exclusively performed on the German manufacturing industry.
Grzybowska and Łupicka, 2017	Survey questionnaire filled by industry experts (20 in total, 10 from each industry interviewed).	Survey results representative of only the automotive and pharmaceutical industries, in Poland.
Wilke and Magenheim, 2017	Attempts to map the learning territory in I.D. 4.0, by employing a multi-strategy approach (structured interviews, consisting of open questions and rating scales).	Survey is limited to just 15 participants (1 female and 14 male), ranging from 16 to 35 years. There is obvious sex bias and age limitation. Majority of the candidates interviewed were trainees in the metal cutting industry. The data is thus, not applicable universally. Finally, the data is limited only to Germany, where the study was conducted.
Bogoviz et al., 2019	Predominantly qualitative and mildly quantitative comparison of I.D. 4.0 competencies and desired skills between developing/developed countries.	I.D. 4.0 indicators are only assessed for four developing economies: India, China, Brazil, and the South African Republic (SAR). These are, in no way at all, representative for all the developing nations. Indicators are not actually measured but obtained from country specific reports. Unlike the developing countries (which are clearly identified for this work), no "developed" country is identified to perform these comparisons. Results and overall conclusions are probabilistic, needing further rigorous assessment.
Neaga, 2019	Identification of E.D. 4.0 core competencies is performed via a systematic literature review and using content and thematic analysis.	The analysis is limited exclusively to university undergraduates, masters, and Ph.D. students in the United Kingdom, predominantly for manufacturing, automotive engineering, and supply chain management programs.
Uhlemann et al., 2019	An extremely exhaustive study, that incorporates both literature review (227 papers, 146 classified as fundamental, and 81 applications oriented) and product design and engineering (PDE) insights, gathered by interviewing 27 PDE experts (both academic and industry experts), across 25 disciplines. Several universities and firms are studied, globally.	The literature review was restricted to the discipline of chemical engineering. This is because the work is on understanding PDE approaches in chemical engineering. We identify this work as a representative example for investigating E.D. 4.0 efforts thoroughly, prior to mapping with I.D. 4.0's ever-changing requirements.
Beke et al., 2020	Questionnaires answered by B.Sc. engineering students at Óbuda University. I.D. 4.0 competencies identified by interviews conducted by the automobile industry, from 2015 to 2020. Responses evaluated using a Pareto diagram.	Analysis limited to engineering students, for one university, in Hungary. I.D. 4.0 skills limited to those locally perceived by the Hungarian automobile industry. Only six companies responded (four international and two medium sized).
Jerman et al., 2020	Audio-taped interviews of 14 subject experts (3 employed at the government, 5 higher education professors, 5 from the automotive industry, and 1 from the chamber of commerce).	Possible sex bias (12 males and 2 females). Results apply to Slovenia. Interestingly enough, the automotive industry keeps getting interviewed predominantly. Experts' views and perceptions likely to vary immensely by sector.
Maisiri and van Dyk, 2020	Literature review from Scopus and Web of Science (WOS), with 25 papers satisfying the inclusion and exclusion criteria. Key competencies characterizing E.D. 4.0 were identified.	Analysis restricted to the domain of Industrial Engineering.
Mian et al., 2020	SWOT (Strength, Weaknesses, Opportunities, Threats) analysis to understand I.D. 4.0 efforts to revamp E.D. 4.0, in sustainability. 200 non-random respondents with engineering education experience (faculty, students, and researchers) were chosen.	22.78% of the survey respondents reported they were unaware about E.D. 4.0. This effectively devalues the efficiency of the non-random sampling methodology employed. None of the survey respondents possessed detailed knowledge about the core components that comprise E.D. 4.0. All survey respondents were also chosen from King Saud University, Saudi Arabia.
Lin and Low, 2021	A case study is performed to assess the alignment of E.D. 4.0 targets with the continuous education training (CET) program, within Singapore.	Results are only applicable to Singapore.
Silva et al., 2021	Mapping attempts to identify advancements toward realizing E.D. 4.0 goals. Out of 1,732 studies, 78 were eventually selected.	There were several databases that the authors investigated: SCOPUS, ACM, IEEEXplore, SBIE, and RBIE. However, the selected studies are only limited from 2015 to 2018. Our study is performed for the last decade, and thus, captures much more details, and is likely a more accurate indicator for E.D. 4.0/I.D. 4.0 mapping.
Costan et al., 2021	Meta-analysis performed for 299 articles; 30 meet the inclusion criteria. 12 barriers to E.D. 4.0 identified for developing countries, compared to the developed counterparts (using the PRISMA statement approach).	Search performed only in the Scopus database. Only published articles from 2015 to 2022 were selected. The subject areas being assessed did not include the humanities (with the exception of one work reviewing Digital English and E.D. 4.0, for I.D. 4.0). Also, over 25% of the papers assessed originate from Malaysia. This may very likely skew the obtained conclusions toward a more localized perspective.
González-Pérez and Ramírez-Montoya, 2022	Systematic literature review in the Scopus and Web of Science (WOS) databases. 56 articles successfully pass the Inclusion and Exclusion criteria.	Search performed for articles published only between 2006 and 2020.
Ramírez-Montoya et al., 2022	Literature review is performed on 48 articles; VOSviewer is used to identify the search keywords, to further refine the papers to 35.	The literature review is performed only on two dabases, namely, Scopus and the Web of Science (WOS). 35 articles effectively met the inclusion and exclusion criteria. The extremely limited number of articles searched (from 2002 to 2021), provide limited insights. We employ VOSviewer across several databases, to identify the defining competencies of E.D. 4.0.

TABLE 2 A chronological summary of E.D. 4.0 competencies as identified by some prior researchers.

References	Key competencies identified
Hecklau et al., 2016	**Technical:** state-of-the-art knowledge, technical skills, process understanding, media skills, coding, understanding IT security. **Methodological:** creativity, entrepreneurial thinking, problem solving, conflict solving, decision making, analytical skills, research skills, efficiency orientation. **Social:** intercultural skills, language skills, communication, networking, teamwork, leadership, knowledge transference, cooperation/compromising ability. **Personal:** flexibility, ambiguity tolerance, learning motivation, compliance, sustainable mindset, ability to work under pressure.
Grzybowska and Łupicka, 2017	Creativity, entrepreneurial thinking, problem solving, conflict resolution, decision making, analytical skills, research skills, efficiency orientation.
Prifti et al., 2017	Communication with people, IT/technology affinity, big data, problem solving, life-long learning, interdisciplinary work environment, network technology, M2M communication, modeling/programming, data/network security, business process management, collaboration, teamwork, decision making, leadership skills, service orientation, creativity, self-management.
Ramirez-Mendoza et al., 2018	Computational skills, virtual collaboration, resilience, social intelligence, novel and adaptive thinking, load cognition management, sense making, new media literacy, design mindset, transdisciplinary approach.
Neaga, 2019	**General:** computation skills, virtual collaboration, resilience, social intelligence, novel and adaptive thinking, load cognition management, sense making, new media literacy, design mindset, transdisciplinary approach. **Disciplinary:** fundamental sciences (maths, physics, chemistry, biology, statistics, and coding), applied sciences (materials, manufacturing, control principles, signal processing, applied statistics, and system engineering), industrial automation and control, production, business and management, advanced manufacturing, information and communication technologies.
Das et al., 2020	Complex problem solving, critical thinking, creativity, people management, coordinating with others, emotional intelligence, judgment and decision making, service orientation, negotiation, cognitive flexibility.
Jerman et al., 2020	Continuous learning, flexibility/adaptation to change, technical literacy, problem solving, soft skills, critical and analytical thinking.
Maisiri and van Dyk, 2020	**Soft skills:** critical thinking, agile problem identification and problem solving, communication skills, open minded thinking. **Knowledge:** programming in R, Scala, Python, and PySpark, coding, big data analytics. **Technical skills:** data analysis, visualization and cleaning, pattern recognition, data corroboration.
Miranda et al., 2021	**Transversal (or soft) skills:** critical thinking, cooperation, collaboration, communication, creativity and innovation. **Disciplinary (or hard) skills:** training and development, research/design/implementation of new strategies, technology-based solutions, emerging best practices.
Ramírez-Montoya et al., 2022	**Disciplinary:** functional, technical, and technological knowledge/skills, research, design, create and implement technologies, emerging technologies, technology-based solutions. **Transversal:** critical thinking, systemic thinking, scientific thinking, innovative thinking.

Any sub-classification of competencies are shown in bold.

TABLE 3 A summary of some key prior literature review/bibliometric analyses, that attempt to understand, classify, and/or capture/map E.D. 4.0 to I.D. 4.0.

References	Number of articles after a first screening	Number of duplicate articles	Number of articles after implementing the inclusion and exclusion criteria
Prifti et al., 2017	26	N/A	17
Da Costa et al., 2019	1,925	547	911
Uhlemann et al., 2019	227	N/A	N/A
Maisiri and van Dyk, 2020	303	34	25
Costan et al., 2021	299	N/A	30
Silva et al., 2021	223	Detected, unreported	78
González-Pérez and Ramírez-Montoya, 2022	113	41	56
Ramírez-Montoya et al., 2022	48	8	35

TABLE 4 Additional elements incorporated into the core components of E.D. 4.0 presented by Miranda et al. (2021) to obtain the combined reference framework.

E.D. 4.0 competencies		Information and communication technologies (ICT)	
Transversal	Disciplinary	Technology-based	Tools and platform
Leadership, strategic view of knowledge, self-organization, feedback, problem-solving, pro-activity, inter-disciplinary, teamwork, initiative, flexibility, adaptability, self-management	Sustainability, automation, information and communication technology	Big data, information and communication technology, neural network and others	Augmented reality, embedded systems, integrated systems

is identified by incorporation of a set of E.D. 4.0 elements identified by Kipper et al. (2021), which are then assimilated into the core components' competencies (transversal and disciplinary) and ICTs (technology-based and tools and platforms) of Miranda et al. (2021), as presented in Table 4. The identified elements in the reference framework are used as keywords or search parameters, when applying the inclusion and exclusion criteria set for our database. A systematic literature search (SLS) is conducted first to analyze these articles, supported by a bibliometric analysis performed using VOSviewer. The software analyzes databases, based on the Visualization of Similarities (VOS) algorithm, as proposed by Van Eck and Waltman (2007), which visualizes both direct and indirect connections between entities, by classifying these relationships into one of three categories: network, overlay, or density (Ejsmont et al., 2020). The strongest clusters typically appear in the center of the generated color-coded plot, and signify a contribution/factor that relates strongly, and more diversely (Waaijer et al., 2011), among the papers analyzed in the database. The exact refinement criteria employed for our review arise from the protocol detailed by Verner et al. (2012), which begins by posing appropriate RQs, executing the search process to obtain an initial number of articles. The articles identified from this first search are further refined through the inclusion and exclusion criteria (to obtain a smaller number of even more relevant articles), and then, selecting and extracting the relevant data. Section "3.1. Hypotheses and research questions" describes these details in more detail, as applied in the context of this work. From this refined dataset, we aim to capture the most updated mapping tendencies of E.D. 4.0 to I.D. 4.0, by formulating a set of RQs, after identifying the key objectives this work seeks to address; these are as follows,

(a) Identify the extent of E.D. 4.0, adopted as a formal framework toward mapping I.D. 4.0,
(b) Identify the extent of adopting E.D. 4.0 components from the "formal" E.D. 4.0 framework: and
(c) Analyze trends toward implementing the E.D. 4.0 framework and its core components, over the last decade.

Having identified our "reference framework," the above three objectives are investigated as follows. Objective (a) is assessed by identifying the generated clusters, weight attributes, and skewness of E.D. 4.0 elements, from the visual network information generated by VOSviewer. Objective (b) is studied by identifying E.D. 4.0 components from our reference framework, which currently exhibit a lack of standardization. As a direct consequence, this exercise results in the identification of novel elements which can ensure a complete, updated mapping framework. Finally, objective (c) is studied by investigating trends in implementing our reference framework over the last decade and identifying major indicators of implementation and/or mapping efforts. It is noted that our proposed three objectives focus on two thematic issues: the standardization of the E.D. 4.0 framework, and its effectiveness and progress made toward implementing it. An additional theme is also recognized as scope for future work, namely, the challenges which exist toward implementing E.D. 4.0 between developed and developing countries. We hypothesize that implementation trends our reference framework might be considerably different across these two socio-economic groups.

To realize our identified objectives, four RQs are formulated, as summarized in **Table 5**. These are as follows:

> *RQ_1: To what extent has E.D. 4.0 been integrated as a framework for mapping with I.D. 4.0?*

This question aims to identify the VOSviewer generated plots, which can provide insights on how much of E.D. 4.0 currently aligns toward I.D. 4.0 targets, and how much mapping work remains to be performed. RQ_1 is primarily aimed at addressing objective (a), and these insights generated from the bibliometric analysis will strategically identify the scope of future mapping endeavors that should be performed.

> *RQ_2: Which E.D. 4.0 components are successfully identified in the literature, out of the composite E.D. 4.0 framework?*

The purpose of this RQ is to identify objective (b) and recognize two important factors: the components which have been successfully mapped out from the E.D. 4.0 literature, and the components that are yet to be mapped out from our reference framework. Together, RQ_1 and RQ_2 are formulated to test our hypotheses about the lack of a standard implementation of E.D. 4.0 framework, to meet I.D. 4.0's ever-evolving requirements. As we will shortly see, this exercise results in a novel, revamped identification of E.D. 4.0 competencies, as I.D. 4.0 demands continue to evolve with time, across an ever more competitive global workspace.

> *RQ_3: How have RQ1 and RQ2 been addressed, over the past decade?*

This RQ is aimed to investigate objective (c), through comparative (literature reviews) and visual (VOSviewer maps) studies. The choice of a decade was considered an appropriate timeline to track the temporal variation in E.D. 4.0 requirements, as a response to I.D. 4.0's ever changing requirements. RQ_3 attempts to investigate and capture these temporal trends regarding E.D. 4.0 research as a unified framework, or present core components of E.D. 4.0 outside of the reference framework. By analyzing the trends obtained from these plots, in the context of RQ_3, we successfully identify the core competencies of E.D. 4.0 today, and notably, they are somewhat different from the competencies identified during the last decade. This response is expected: as societies evolve and job requirements become even more rigorous and demanding, in an increasingly globalized, AI-driven world, E.D. 4.0 expectations also evolve to match the need to educate adaptive and industrially competent students, over current and future generations. It must be emphasized that while analyzing RQ_1-RQ_3, the authors have ensured that the data collected is impartial, and applicable globally (not restricted by geographical, linguistic, and or cultural barriers). Consequently, the conclusions obtained from this work are far more universal and applicable across almost every sector today.

> *RQ_4: How do E.D. 4.0 efforts compare between developing and developed countries?*

The purpose of RQ_4 is directed more toward identifying the scope of future research work in this field, following our bibliometric analysis. It appears evident that the extent to which E.D. 4.0 efforts would be implemented between developed and developing countries, because of (likely) economic, social, political, and systemic barriers. Also, RQ_4 attempts to identify the key factors responsible for the gap in adopting the E.D. 4.0 framework between nations, thereby acknowledging a currently existent imbalance of the E.D. 4.0 framework. This should not be perceived negatively,

TABLE 5 Themes and research questions formulated to test the objectives, in the reference framework.

Themes	Research questions (RQ)
Standardization of the E.D. 4.0 framework	RQ_1: To what extent has E.D. 4.0 been integrated as a framework for mapping with I.D. 4.0?
	RQ_2: Which E.D. 4.0 components are successfully identified in the literature, out of the composite E.D. 4.0 framework?
Effectiveness and progress in implementing E.D. 4.0	RQ_3: How have RQ1 and RQ2 been addressed, over the past decade?
Challenges toward implementing E.D. 4.0	RQ_4: How do E.D. 4.0 efforts compare between developing and developed countries?

but instead, should serve as a motivator for future research arising out of the framework detailed out in this manuscript.

3.2. Search criteria

To ensure a perfectly unbiased treatment of the search procedure, the search criteria employed several electronic databases, namely, Scopus, Science Direct, and IEEE, which are known to be online repositories for articles published in journals, books, and articles. The search period was between January and June of 2022. The search strings for RQ1 are chosen to be "Education 4.0" AND/OR "Industry 4.0" AND/OR "Fourth industrial revolution," which led to 384 articles being matched with the search string. These articles are open access, peer-reviewed in journals, conference proceedings, across all languages, and almost 50% of the literature applies to the subject field of "Engineering and Computer Science," as shown in **Figure 1**. This is not surprising at all, since this field has seen a tremendous boom over the last decade. Rather, this is reassuring, because it successfully demonstrates that the first filter for our search criteria works well. From this initial dataset, and to obtain a more representative idea about the exact trends that RQ_2 seeks to identify, E.D. 4.0 core components were searched exclusively in the domain of "Engineering Education," where substantial studies have been performed traditionally (in fact, most of the literature cited in this work also happens to be from this field). As is evident from **Figure 1**, "Engineering and Computer Science" represents 50% of the articles found in the databases, followed by "Social Sciences" at 24%, "Decision Sciences" at 8%, "Life Sciences & Medicine" at 5%, "Energy" at 3%, and "Business, Management and Accounting at 6%" and "Art and Humanities" at 2%.

3.2.1. Inclusion and exclusion criteria

Inclusion and exclusion terms were manually selected from the search list provided by VOSviewer to align found words with all prior identified core competencies. Search words were used as basic search terms.

3.2.2. Data selection and extraction

Records were initially identified through database searching engines, followed by a manual screening process to eliminate duplicates. Text articles were then assessed for title and abstract matches.

3.2.3. Data synthesis

A rigorous bibliometric analysis is conducted via VOSviewer, which allows the creation of maps, based on analyzing the network data of various scientific publications and journals. The maps are created by using bibliographic databases obtained from Scopus, Science Direct, and the IEEE databases. Network maps include items or objects of interest between any pair of items, with a detectable link or connection. Each link is assigned a strength (a positive numerical value); the higher the strength, the stronger the link between the items. Items are also grouped into clusters and may have various attributes, for example, cluster numbers for example. Likewise, weight attributes indicate the relative importance of an item, and an item with higher weight is therefore more important than its lower counterpart. In the network visualization of these maps, higher weight items feature more prominently. To support the discussion of RQ3, supplementary graphs were created in Microsoft Excel, including trends and pie charts. While bibliometric analyses provide some immediate clear advantages such as the quick assessment of research impact and scalability of large volumes of data, additional metrics may be necessitated to distinguish them better, and perform more in-depth studies. We also recognize that bibliometric analyses may even possibly skew the research toward the most cited contents, thus, bibliometric indicators should be treated more as a first filter, which would then serve as a starting point for more detailed pedagogical research. Nevertheless, bibliometric analyses are an extremely practical visualization tool which can successfully cluster large volumes of research data; and further analyses are likely needed to obtain more conclusive insights.

4. Results

In this section, the findings arising from the bibliometric analysis implemented by us are summarized, to answer and add more context to the initially posed RQs. We also consciously present the discussion and identify key strategies that may be employed in the future to map E.D. 4.0 to I.D. 4.0 more comprehensively, considering the gaps in the process, as identified from our literature review.

> *RQ1:* To what extent has Education 4.0 (E.D. 4.0) been fully integrated as a framework for mapping with Industry 4.0 (I.D. 4.0)?

Figure 2 depicts the network map, comprising of several clustered items, related to the posed RQ RQ_1. VOSviewer generates six clusters, around highly weighted items, such as performance (virtual reality, simulation, video, and ICT), the internet (IoT, cloud computing, and education system), methodology (flexibility, creativity, and soft skill), quality (blended learning and e-learning), stakeholder (complex thinking and critical thinking), and curriculum (cyber-physical system and

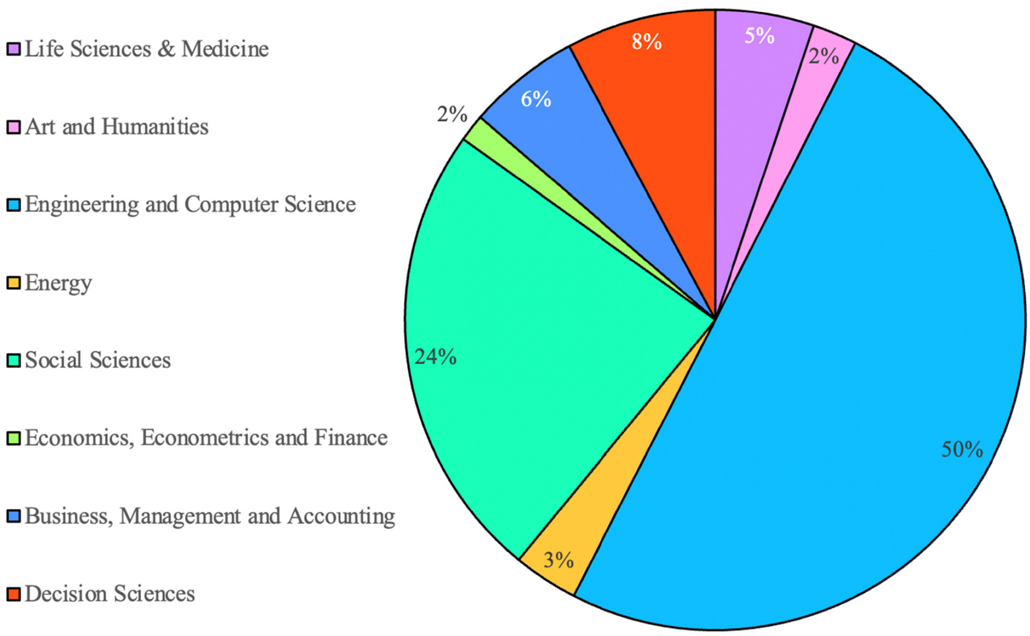

FIGURE 1
The distribution of E.D. 4.0 articles, according to various subject domains, as of 2022.

augmented reality). There is clear evidence of some core components fitting this mapping model, when searching by the keyword "Education 4.0." However, some currently missing elements remain yet to be mapped, such as hybrid-based learning, active learning, problem-based learning, technology-based items (ML, big data, data science and data analytics), AI, and infrastructural setups. Our bibliometric analysis also appears to signal toward new emerging pathways (these are the resulting clusters on the map), that characterizes E.D. 4.0 currently; the major ones being methodology, performance, mathematics, quality, internet, video, and AI, and some minor ones being attitude, cloud computing, blended learning, experiment, creativity, sustainability, initiative, complex thinking, IoT, education system, simulation, virtual reality, etc.

> *RQ_2:* Which Education 4.0 components are identified in the literature, out of the entire Education 4.0 framework?

RQ_2 builds on the key competencies obtained from RQ_1, as specifically applicable within the context of the field of "engineering education," since "Engineering and Computer Science" represents 50% of the articles that were found to answer RQ_1. **Figure 3** presents the network map generated by VOSviewer (2022), identifying clustered items related to RQ_2, by analyzing over 1965 articles (open access, peer-reviewed journals, conference proceedings, and across all languages) found in our databases. VOSviewer generates eight clusters around highly weighted items such as motivation (soft skill, teaching material, engineering curriculum, sustainable development, simulation, collaborative learning, critical thinking, experiential learning, and blended learning), the IoT, modeling (practical implication and technological innovation), cyber-physical systems, big data (cloud computing and data science), algorithms (neural networks and cloud), ML and AI. The links and weights of the clustered items in the mapping represent all core components that E.D. 4.0 comprises of. The articles analyzed in our datasets are over 35 times more than those of Maisiri and van Dyk (2020), over 3 times more than Kipper et al. (2021). It is therefore natural to conclude that the clusters identified from our bibliometric analysis is likely to be much more representative of current E.D. 4.0 competencies, as an immensely larger dataset was fed to VOSviewer for analysis. Simultaneously, our results are unlikely to be restricted by geography and are far more globally applicable, unlike almost all previous work in this field. We observe that there are in fact, several "core" components that comprise the current day perception of E.D. 4.0, and we postulate that these clusters are likely to increase even more in the future, as I.D. 4.0 requirements will likely become even more stringent, across a more-competitive workspace.

The bibliometric analysis provides us with two trackable parameters for each identifiable trends of E.D. 4.0 – the link strength, and the occurrence. While the link strength is a measure of the number of publications where a keyword occurs, the number of times it occurs among the publications identified by the link strength detection criterion, is the occurrence. **Figure 4** shows the link strength matches and values arising from our bibliometric analysis; as expected, Industry 4.0 and engineering education are the two keywords that record the largest values, at 100 and 93, respectively. In fact, what these numbers suggest is that almost all publications that study I.D. 4.0, also investigates its relationship with E.D. 4.0. To comprehensively understand the relative characteristics of E.D. 4.0 that are currently perceived as valuable by I.D. 4.0, these two search keywords are removed, and all remaining keywords arising from the bibliometric analysis are analyzed in **Figure 5**. The trends are most interesting and reveal some notable shift: ML, the IoTs, and AI emerge as the top three contributors to the currently "perceived" definition of E.D. 4.0. This

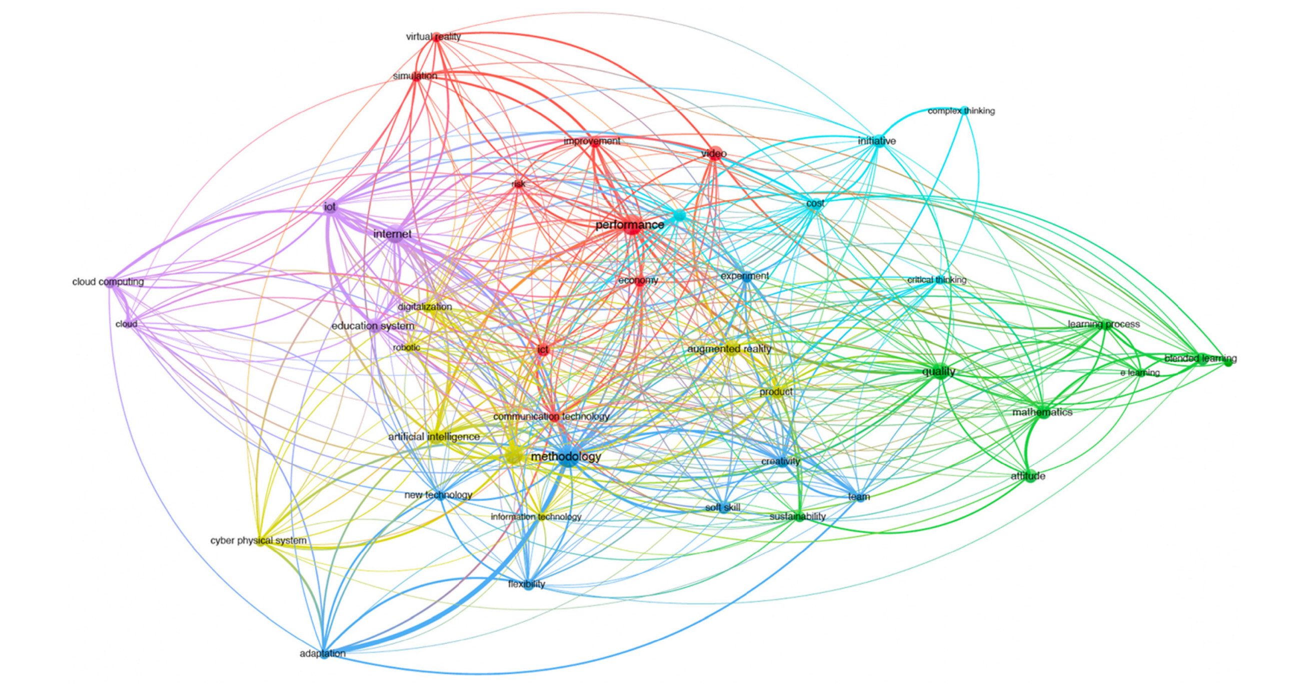

FIGURE 2
The network map obtained as a response for RQ1, which identifies several research clusters which characterize strategies implemented to map E.D. 4.0 competencies to I.D. 4.0 needs.

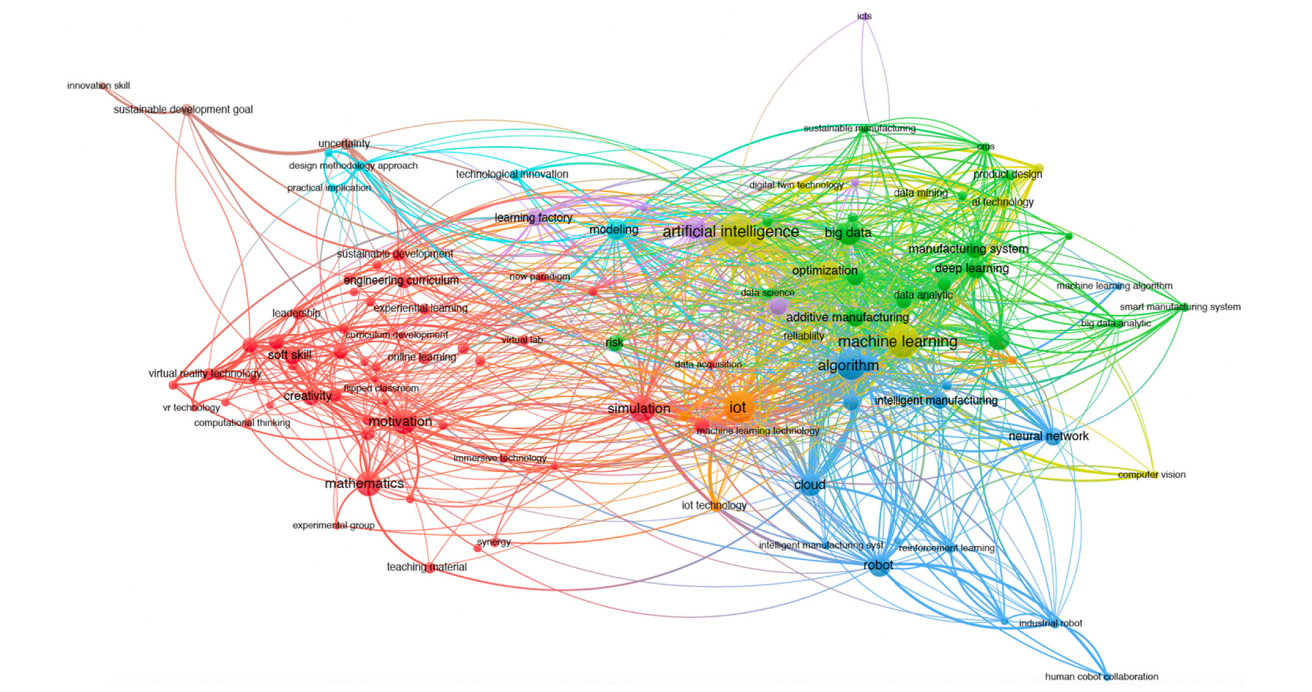

FIGURE 3
The network map obtained as a response to RQ2, identifying the core components of E.D. 4.0.

is a stark difference, compared to previous works in the literature, and it appears that a global transformation is occurring. I.D. 4.0's expectations toward E.D. 4.0 indicates a predominantly technology driven transformation. What is most notable is that almost all the clusters identified are technology-based, and it appears that soft skills-based competencies are perceived less valuable, as the global economy becomes more digitized, and technology-driven. Figure 6 groups these keywords by relevance index into the four

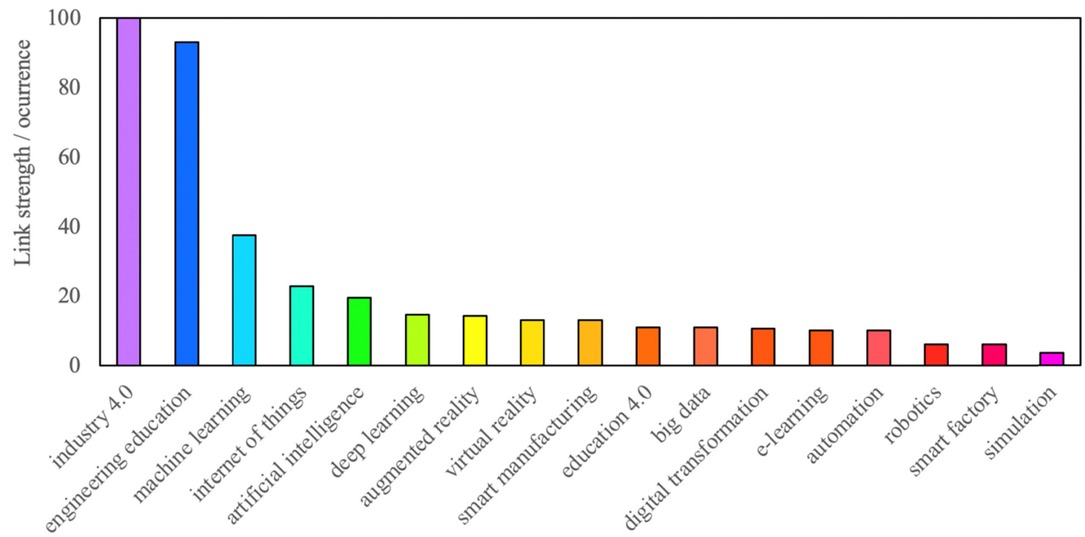

FIGURE 4
Link strength values for matching elements between I.D. 4.0 and E.D. 4.0.

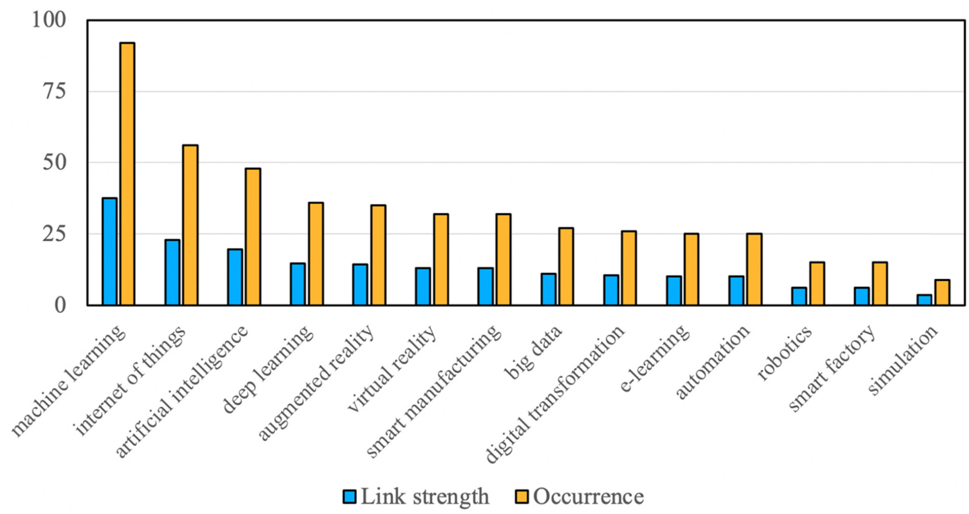

FIGURE 5
Link strength and occurrence values for the identified clusters arising from our bibliometric analysis.

core components of E.D. 4.0; this index is estimated as the average total link strength and the number of occurrences reported by VOSviewer. The results are extremely enlightening: 70% of research in Engineering Education aims to tackle to core component of ICT, followed by transversal competencies (11.6%), learning methods-methods (10.7%), ICT based tools/platforms (4.6%) and others, grouping the rest of components (3.1%). We also note that Competencies (14.2%, as a sum of the Transversal and Disciplinary contributions) and ICTs (74.3%, as the sum of Technology-based tools/platforms) are components that remain fully aligned with the current essential competencies (Miranda et al., 2021), and trends driving industry growth (Hernandez-de-Menendez et al., 2020). The major conclusion that arises is that the traditionally identified components and competencies of E.D. 4.0 do not apply as much, and a new definition of E.D. 4.0 must be proposed, keeping in tandem with the current trends and expectations of I.D. 4.0.

> *RQ_3*: How have RQ_1 and RQ_2 have been answered over the past 5 years?

Figures 7A, B present publication trends of articles published from years 2017 to 2021 pertaining to "Education 4.0" (RQ_1) and "Engineering Education" (RQ_2). Trends show a fairly robust increment of publications over the last decade, with the exception of "Engineering Education" during 2020–2021, which show no significant difference in the number of publications. These trends are extremely revealing, as it confirms a rapid increase in interest within the academic community, to comprehensively understand, characterize, and map the structural and functional components

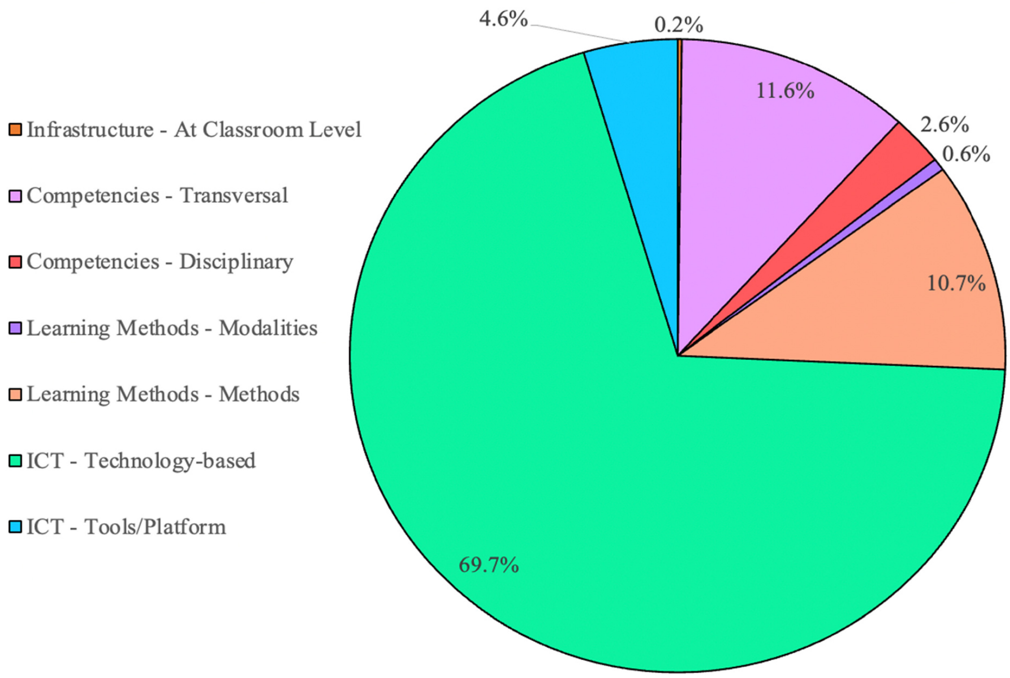

FIGURE 6
Core components of E.D. 4.0, identified by relevance.

of E.D. 4.0 and I.D. 4.0. This trend is very promising and forms a very firm basis for the motivation of this work. Through the methodology and analysis procedure formulated by us here, we are hopeful that this protocol will be adopted by future researchers, to successfully map E.D. 4.0 competencies to I.D. 4.0 demands, as both sectors continue to evolve with time.

> *RQ*$_4$: How do E.D. 4.0 efforts compare between developing and developed countries?

It is evident that there are socio-economic differences between developing and developed nations, and these were especially exacerbated during the COVID-19 pandemic (Gajdzik et al., 2020; Perry et al., 2021). Our bibliometric analysis, which centers around search words such as "Education 4.0," "Education Engineering" and associated E.D. 4.0 core components, reveal that these research efforts are predominantly conducted in developed countries, as compared to their developing counterparts (63 and 51.2% of published articles, respectively). In our classification, we use the conventional definition of developed countries, as those nations which have a Human Development Index (HDI) equal to, or exceeding 0.8 (United Nations, 2022a,b). **Figure 8** shows the most recent values of the HDI, for all countries. The countries which score highest in this criterion are the US, Canada, Australia, the United Kingdom, Ireland, Japan, South Korea, and countries typically comprising the European Union (Austria, France, Sweden, Germany, Finland, Norway, Denmark, Iceland, Spain, and Italy). Trailing slightly lower are some other countries, some of whom are known to have oil export economies (Russia, Kazakhstan, Saudi Arabia, Turkey, Argentina, Chile, Portugal, Croatia, Romania, Bulgaria, etc.). The shift of the world toward online, technocentric learning environment toward E.D. 4.0, arising as a direct result from our bibliometric analysis, is presented in **Figure 9**. A rather different perspective emerges, and three countries (the US, China, and Australia) emerge as locations where this shift has occurred the fastest. A surprising contender is India, which, despite still being a third-world economy, has adapted seamlessly toward a techno-centric economy, thereby also influencing a shift toward E.D. 4.0 measures within the country. But this shift is also somewhat expected, since a significant proportion of the world's programming and information technology (IT) needs are exported to employees in India. After India, some other first-world nations (the United Kingdom, Spain, Germany, Canada, Russia, and Finland) follow. Indonesia also ranks in this list (and this means that the country is moving toward the successful implementation of E.D. 4.0 competencies as identified from our bibliometric analysis), despite not being traditionally considered as a first-world nation. Predominantly, we observe research efforts in developed nations to mainly tackle the ICT core component (71.6%), as against the Competencies counterpart (52.1%) (ITU, 2022).

When **Figures 8**, **9** are compared, an interesting conclusion arises: some first-world nations are yet to revamp their pedagogical and university-based education to meet the demands of I.D. 4.0. As the world becomes increasingly more digitized, we identify that several first-world economies hold immense potential to train future skilled workers who can address the current global I.D. 4.0 demand shortage – some of these being Canada, France, Brazil, South Africa, Japan, Italy, Sweden, Norway, Turkey, Saudi Arabia, and Mexico. We clarify here that our analysis is limited by the available data online, and as countries change their global positioning and strategic outlook toward the E.D. 4.0 metrics, as

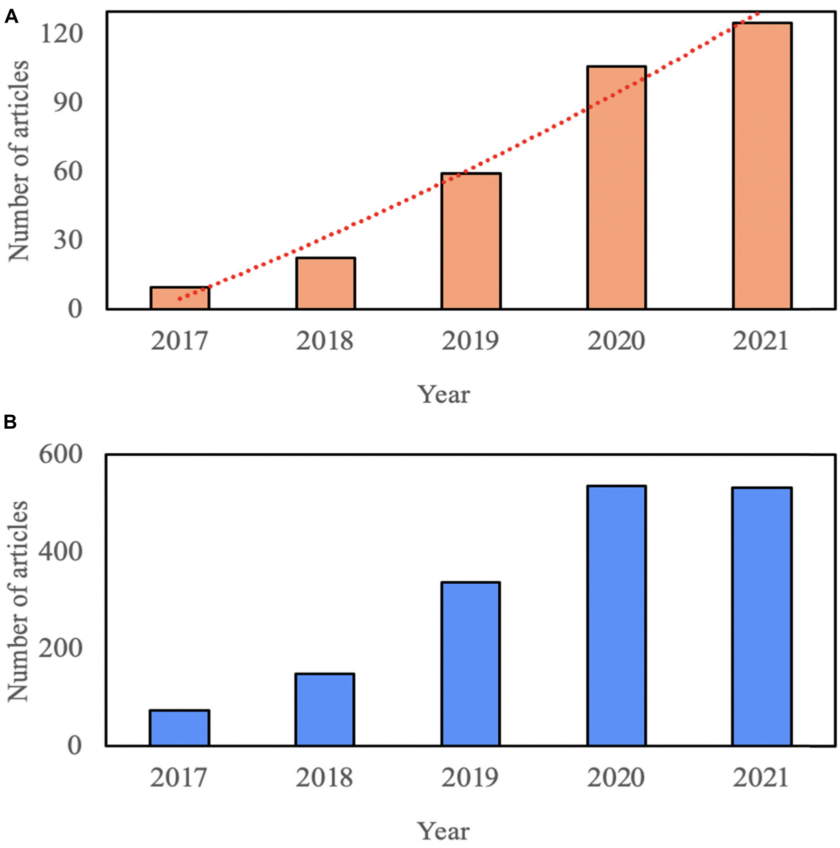

FIGURE 7
Trends of articles published **(A)** in E.D. 4.0 and **(B)** in Engineering Education and E.D. 4.0.

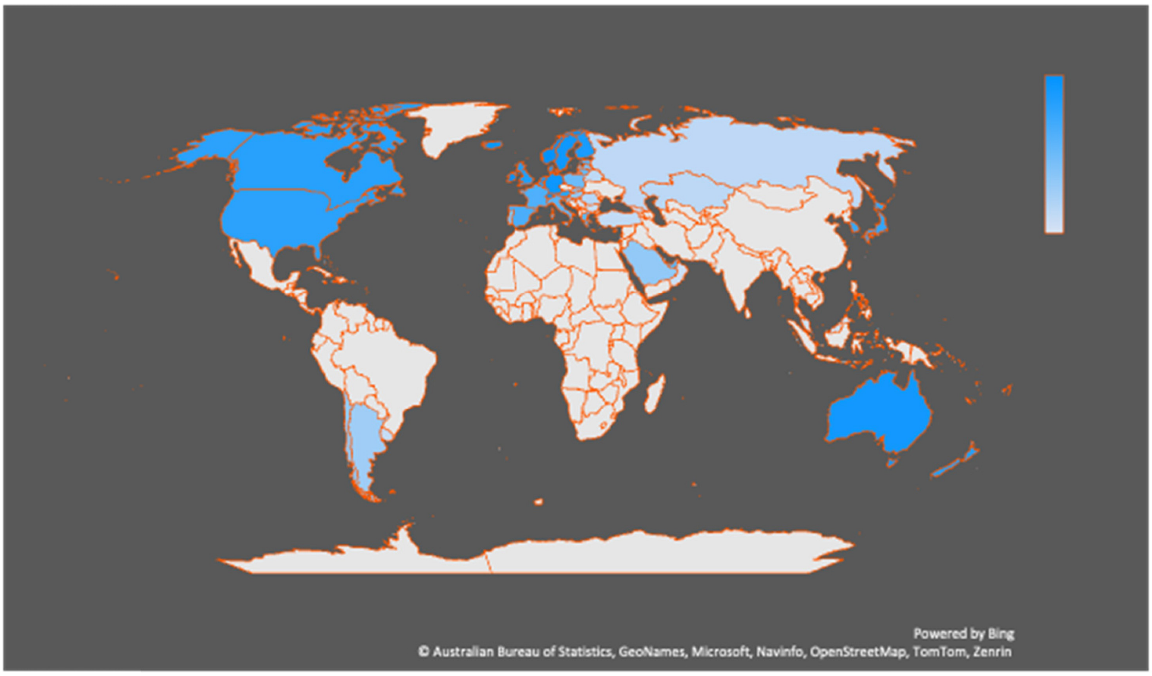

FIGURE 8
Global distribution of the Human Development Index (HDI), as of 2022. As expected, the HDI is stronger for developed economies.

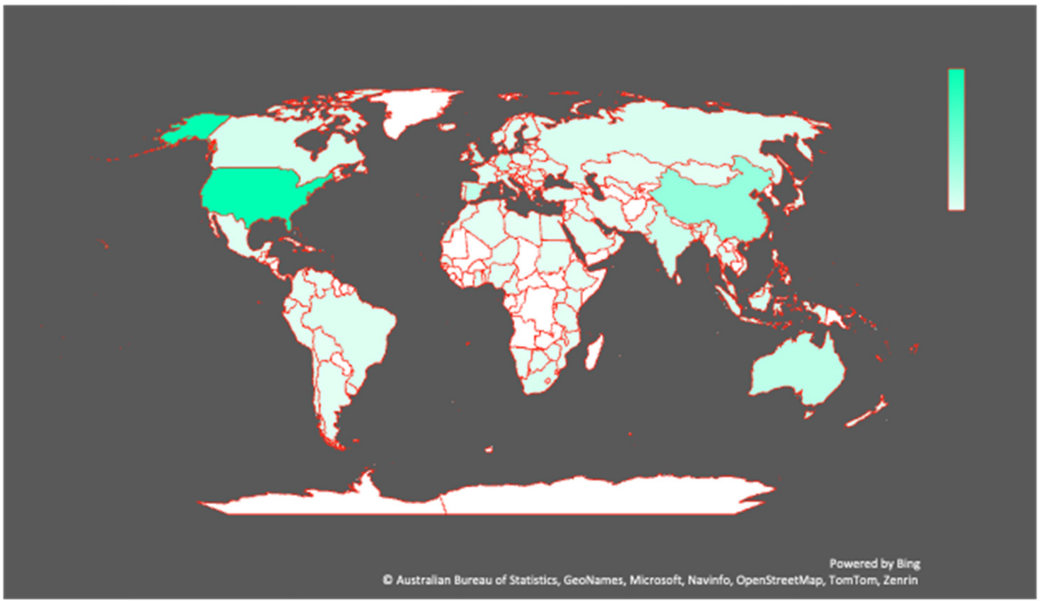

FIGURE 9
Global map representing efforts by countries toward the adaptation of online learning-based E.D. 4.0 environment.

identified by our analysis, these conclusions will evolve over time. It may be worthwhile to perform separate bibliometric analyses within each country and compare which sectors of E.D. 4.0 feature prominently across which states/provinces. Very likely, national economics and the overall industrial asset mix will dictate which aspect of E.D. 4.0 is valued by a nation, when training a generation of students to match its I.D. 4.0 demands.

5. Discussion

The network map that seeks to answer RQ_1 (**Figure 2**) show that defined clusters do not necessarily match core E.D. 4.0 components, although its elements are clearly found when linked to these clusters. This graphically strengthens the evidence toward implementation of E.D. 4.0 elements, but no formal E.D. 4.0 framework to map with I.D. 4.0 exist currently. The unbalanced weight distribution in the clusters, show heavier contributions for the clusters "performance," "methodology," and "internet," overshadowing other critical elements included in the groups quality and curriculum. Moreover, the identified missing elements do not necessarily relate to the lack of implementation of E.D. 4.0 core components but instead occur due to not adopting a standardized E.D. 4.0 definition/framework.

Further evidence of this claim is obtained from the network map that aims to answer RQ_2 (**Figure 3**), where no universally adopted educational framework appears to fully encapsulate and capture I.D. 4.0 needs; however, efforts to match I.D. 4.0 requirements with E.D. 4.0 training which can suitably prepare the workforce in higher education institutions have been massively adopted. Both maps in **Figures 2**, **3** reveal that these efforts might result in discrepancies when assessing the effectiveness of implementing core components, item weight imbalances, potential duplication of elements, and mismatch with I.D. 4.0 priorities. Therefore, it is recommended to periodically update the reference framework to capture the evolution of the I.D. 4.0 requirements; task must be consensually defined under a standardization process. However, a summarization of the major weighted clusters (which are the closest match to I.D. 4.0 requirements), may form a definition for E.D. 4.0 (also summarized in **Figure 10**).

5.1. Definition

Education 4.0 is an educational framework that strategically incorporates competencies such as mathematics, modeling, AI, simulation, ML, the IoT, deep learning, big data, neural network, manufacturing system, robotics, motivation, cloud, etc. into the learning experience, to match the current requirements of I.D. 4.0.

It must be immediately emphasized that this definition is not cast in stone, is by no means exhaustive, and is likely to evolve with time, depending on I.D. 4.0 demands. As of today, this current definition of E.D. 4.0 appears to be extremely reasonable (even to the lay observer), and in tandem with global industrial trends which favor a move toward a more digitized, computer-based economy and workspace. This is also one of the key tenets of the vision of E.D. 4.0, as detailed by Fisk (2017) who imagines a future scenario where "man and machine align to create new possibilities." Therefore, we have demonstrated that bibliometric analyses may be employed to arrive at a (hopefully) universally accepted definition of E.D. 4.0, which will of course need periodic recalibration/updates (say every 5 years, or perhaps a decade). Also, bibliometric analyses can prove useful to inform the pedagogical process across universities and institutions of higher learning, to revamp course curricula in accordance with ever-evolving new market demands/trends, which directly translate to shifts in I.D. 4.0 trends.

It must be immediately emphasized that this definition is not cast in stone, is by no means exhaustive, and is likely to evolve

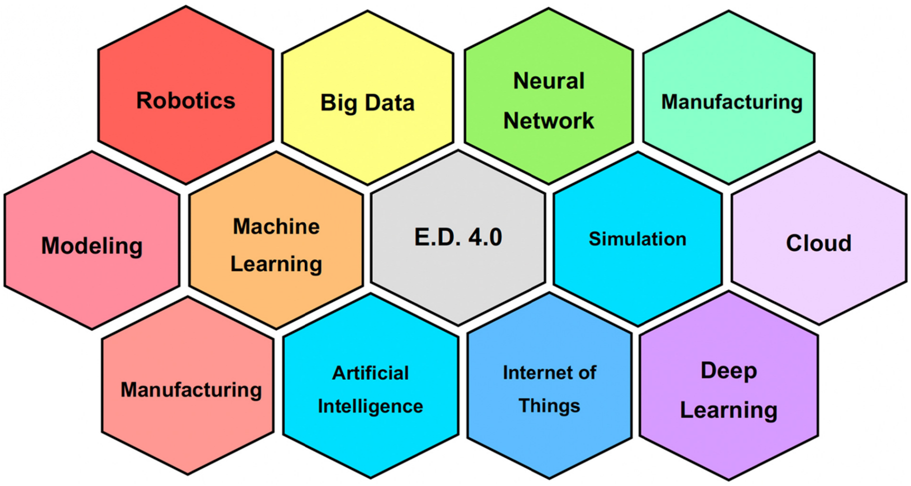

FIGURE 10
Current components of E.D. 4.0, as identified by bibliometric analysis. These elements form the most updated markers of E.D. 4.0 and align closest with the current requirements of I.D. 4.0.

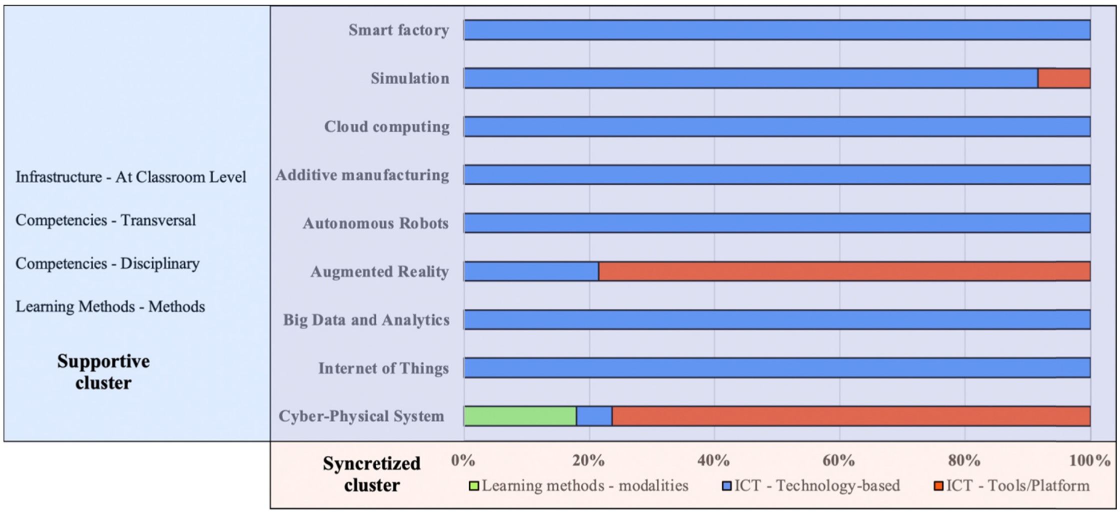

FIGURE 11
Syncretized and supportive clusters, identified post bibliometric analysis, to map I.D. 4.0 requirements, to E.D. 4.0 core elements.

with time, depending on I.D. 4.0 demands. As of today, this current definition of E.D. 4.0 appears to be extremely reasonable (even to the lay observer), and in tandem with global industrial trends which favor a move toward a more digitized, computer-based economy and workspace. This is also one of the key tenets of the vision of E.D. 4.0, as detailed by Fisk (2017) who imagines a future scenario where "man and machine align to create new possibilities." Therefore, we have demonstrated that bibliometric analyses may be employed to arrive at a (hopefully) universally accepted definition of E.D. 4.0, which will of course need periodic recalibration/updates (say every 5 years, or perhaps a decade). Also, bibliometric analyses can prove useful to inform the pedagogical process across universities and institutions of higher learning, to revamp course curricula in accordance with ever-evolving new market demands/trends, which directly translate to shifts in I.D. 4.0 trends.

To answer RQ_3, we refer to **Figures 7A, B**, which capture the growth trends over the last few years, highlighting the increasing interest and support from researchers/educational institutions in preparing a skilled workforce that can not only meet, but also exceed I.D. 4.0 expectations. This may either be formally stated as E.D. 4.0 efforts, or by conducting research on these core components, as has been the norm within engineering education. An in-depth analysis of the context and context of the significant increase in publications in 2020 and 2021 in **Figure 7A**, shows that the COVID-19 pandemic accelerated the implementation of ICT elements (for instance, meeting platforms like Zoom and Microsoft Teams emerged as major global players in this sector, and continue to do so even now), as these elements allowed teachers and students to minimize educational disruption, while parallelly tuning teaching-learning strategies and hybrid pedagogical methods. Researchers are already on their way to

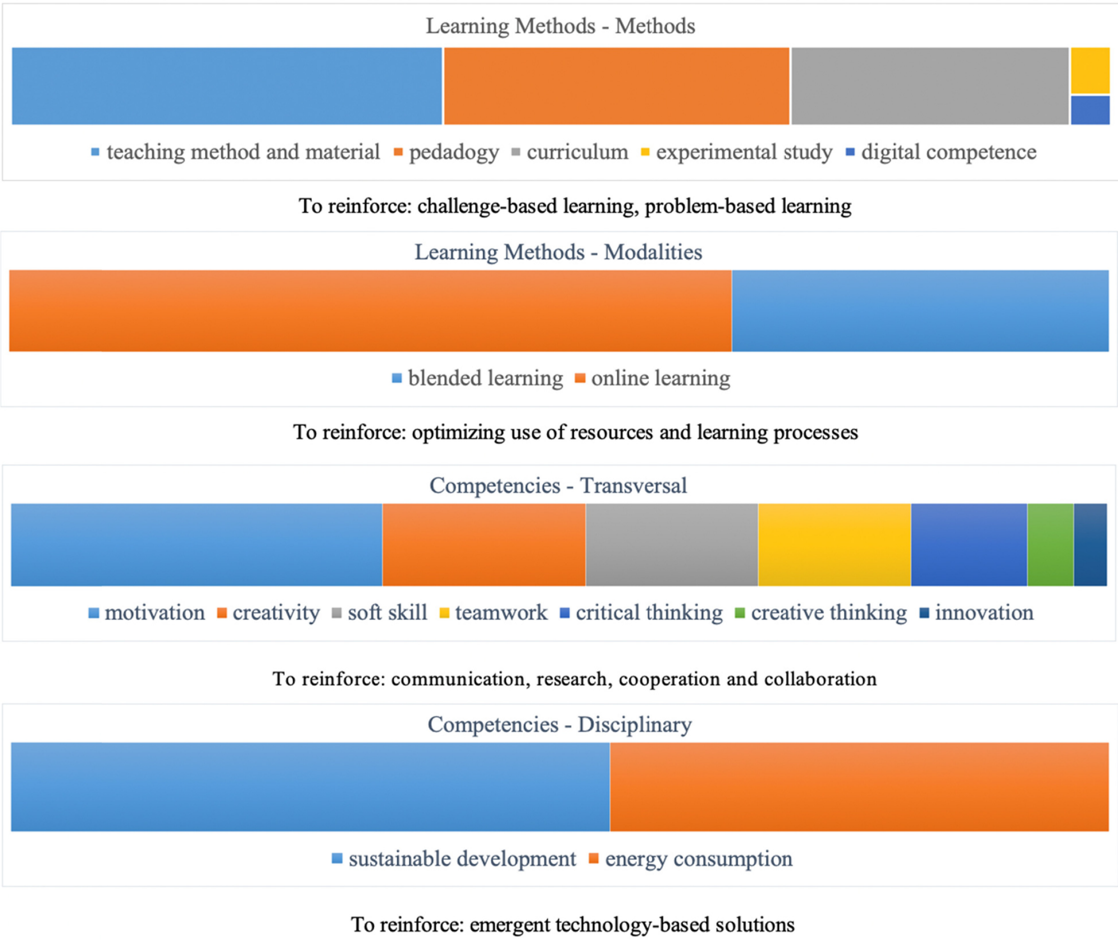

FIGURE 12
Recommendations arising from classification and analysis of E.D. 4.0 clusters. Each of these bars represent graphically the relative abundance/deficit of contributory components toward a specific competency.

evaluating the impact(s) of this paradigm shift on the education industry (Oleksiyenko, 2021; Reimers and Marmolejo, 2022). At first sight, the impact of COVID-19 cannot be observed in the overall publication trends shown in **Figures 7A, B**. However, a restricted search in the identified database, by adding specific keywords such as e-learning, blended learning methods, digital tools, virtual labs, and online assessments, clusters them into one distinct group, which increased 76.2% of components researched between the years 2020 and 2021, as compared to the whole decade. This sudden jump was the response of pedagogical researchers, following the COVID-19 pandemic. The prevalent components fit mainly into the category of Learning Methods in Modalities.

When answering RQ_4, the inequality evidenced in the trends of E.D. 4.0-related published articles might be related to higher ICT indexes in developed countries compared to developing countries, facilitating the implementation of technology-based solutions, and using tools and platforms to support learning modalities and methods. At a regional level, Europe appears to lead research efforts in E.D. 4.0 and Engineering Education per capita (176 and 1313, respectively), as measured as the number of related publications per billion, followed by Oceania (114 and 982), North America (105 and 615), South America (47 and 152), Asia (39 and 143), and Africa (9 and 54). Moreover, 26.9% of all the articles were published under Open Access (OA) modality, and 64.5% of these OA articles were published in developed countries. The number of published research articles might not be the sole evidence of the inequality in implementing E.D. 4.0 between developed and developing countries, but it correlates with the findings of other researchers on evaluating barriers and challenges that impede the universalization of this framework (Bogoviz et al., 2019; Costan et al., 2021; Perry et al., 2021; Wakamo, 2022). Hence, our research results also support increasing efforts to identify, analyze, and tackle systematic barriers toward implementing I.D. 4.0 – E.D. 4.0 in developing countries to reduce education and skills-based inequality. Our work promises to serve as a good reference not only in Europe (which appears to lead in E.D. 4.0 measures), but also for several other continents/countries, owing to the generic nature of the analysis, and its global applicability, simplicity of analysis, and reproducibility. Most importantly, the practical implication of these results will lead to an immediate streamlining of the engineering education research, while providing a robust, comprehensive, reproducible analysis procedure that may be successfully employed toward further global studies. Such a framework will also prove valuable for researchers to compare their pedagogical observations.

Currently, there is no universally agreed upon framework, thus, comparisons are at best only qualitative, and no quantitative comparisons are possible: our procedure overcomes this limitation and incorporates both comparison methods. Additionally, the obtained competencies of E.D. 4.0 are likely to be consistent with industry trends: for instance, today's global workspace seems to be strongly dominated by AI/ML, and these form a strong component of the E.D. 4.0 definition. The promise of a successful mapping between E.D. 4.0 and I.D. 4.0 through our procedure is apparent from the fact that the most influential industry trends, as observed today, form part of the bibliometric analysis inferred definition of E.D. 4.0, which may be interpreted as a successful attempt toward "defining" E.D. 4.0, as interpreted and valued by industries today.

From these in-depth discussions, a global perspective of E.D. 4.0 emerges, as applicable to today's word. Bibliometric analysis also enables us to classify our identified E.D. 4.0 competencies, into two clusters – syncretic and supportive (as shown in **Figure 11**). What we observe is that most of the identified cluster components are predominantly influenced by ICT based technologies, and related tools and platforms (as expected). The relative "mix" of each of these supportive cluster competencies are summarized in **Figure 12**, and additional major insight(s) are gathered.

For *Learning Methods – Methods:* There appears to be a conscious shift toward a revamp of curricula and conscious pedagogical studies. However, experimental study and digital competence score lower; this means that future methods should aim to foster challenge-based and problem-based learning strategies, in additional to existent pedagogical methods, to provide students with a holistic learning experience, while simultaneously empowering them to meet I.D. 4.0 demands in the future.

- For *Learning Methods – Modalities:* There is a slight advantage that online-learning has, compared to blended learning. This trend is in complete accordance with the E.D. 4.0 clusters identified by our bibliometric analysis. It appears that this aspect is almost balanced, and the only recommendation is to optimize the use of proper resources, to facilitate the learning experience for students.
- For *Competencies – Transversal:* There are several contributory components identified – motivation, creativity, soft skills, teamwork, critical thinking, creative thinking, and innovation. First, it is worth observing that creative thinking has taken a backseat, with the advent of a more digitized global workspace. This should be regarded as a global concern, and future pedagogical research should be strategically implemented to foster this extremely crucial skill among future students. Collaborative skills also score lower (a natural consequence of the individual, work-from-home, online workspace that the COVID-19 pandemic ushered in), and students must be also taught the values of communication, research, cooperation, and collaboration in a world that is becoming increasingly individualistic, when it comes to working style. These are important areas for pedagogical researchers to base their future works on.
- For *Competencies – Disciplinary:* There is an almost balanced perspective of energy consumption and sustainable development. However, with the emergence of greener technologies (blue/green hydrogen, biofuels, bio-refineries, etc.), we recommend incorporating emergent (green) technology-based solutions into future course curriculum.

6. Conclusion

The lack of standardization toward defining an E.D. 4.0 framework, for mapping I.D. 4.0 requirements continues to remain an extremely important issue, as has been reiterated in the literature. The conclusions arising from the work of Silva et al. (2021) reveal several initiatives that characterize an alignment of E.D. 4.0, to meet I.D. 4.0 demands. Such initiatives tend to (a) protagonize the student, (b) incentivize active learning, (c) propose the development of practical initiatives, (d) develop skills that are relevant in the 21st century, and (e) enable experiences with emerging computationally aligned resources/processes. Such efforts should be strategically targeted to furnish students with educational skills and assets that are more aligned to I.D. 4.0 needs, to enable maximum chances of employability.

In this work, we propose a method to define E.D. 4.0, through a reverse-engineering of the problem. Rather than attempting to describe E.D. 4.0 competencies from scratch, a bibliometric analysis on the relevant literature provides us clusters, which then form the basis for our definition of E.D. 4.0. Our approach of identifying an integrated reference framework as a source for identifying these E.D. 4.0 elements (clustered across four components) by merging the most up-to-date efforts of Kipper et al. (2021) and Miranda et al. (2021) is non-discriminatory. Over time, as I.D. 4.0 competencies evolve, the same exercise may be repeated after identifying suitable previous reference frameworks, to obtain a more representative definition of E.D. 4.0, as and when required. Verification of the constituent components of the reference framework is implemented using a detailed bibliometric analysis spanning over 1,965 articles, which provides graphical indicators on mapping E.D. 4.0 skills to I.D. 4.0 demands. The advantages of such a bibliometric approach make our analysis global (results are not geographically restricted), and simultaneously, provide a measure of the effectiveness of implementing core components, as VOSviewer can efficiently analyze large volumes of data and provide insight on implementation trends, under minimal time. A detailed analysis of the graphs generated reveals that some of the generated element clusters fall outside the proposed framework, suggesting that the definition of E.D. 4.0 must be expanded to incorporate these clusters as well. Evidence of these additional E.D. 4.0 clusters arising out of bibliometric analyses, outside of any currently accepted formal (and somewhat more theoretical) E.D. 4.0 framework, proves the lack of standardization currently toward defining an exhaustive E.D. 4.0 framework. Thus, further research toward identifying and comprehensively defining these new clusters is needed, otherwise a mapping with the requirements of I.D. 4.0 is likely to remain imbalanced. The extent of universal adoption of the currently existent E.D. 4.0 framework is also assessed, and a clear gap is identified between developed and developing countries, in terms of successful adoption of this framework currently. Our conclusion, which independently arises out of an exhaustive bibliometric analysis, quantitatively supports the (more) qualitative trends observed by prior researchers. In summary, the integration of bibliometric analyses tools toward comprehensively identifying a global definition of E.D. 4.0 is a paradigm shift in the field and is the first holistic technique to capture all elements. Once a robust definition of E.D. 4.0 is formulated, and receives universal consensus, future academics can then map out E.D. 4.0 to I.D. 4.0's requirements.

Data availability statement

The original contributions presented in this study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding author.

Author contributions

SC: conceptualization, methodology, results, analysis, and writing. YG-T and JM: conceptualization, analysis, and proofreading. DG: conceptualization, analysis, writing, proofreading, and principal investigator. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2023.1150190/full#supplementary-material

References

Agrawal, S., Sharma, N., and Bhatnagar, S. (2021). "Education 4.0 to Industry 4.0 Vision: Current trends and overview," in *Proceedings of ICEM 2020: recent advances in smart manufacturing and materials* (Berlin: Springer Singapore), 475–485. doi: 10.1007/978-981-16-3033-0_45

AnTosz, K. (2018). Maintenance–identification and analysis of the competency gap. *Eksploatacja Niezawodność* 20, 484–494. doi: 10.17531/ein.2018.3.19

Bauer, W., Hämmerle, M., Schlund, S., and Vocke, C. (2015). Transforming to a hyper-connected society and economy–towards an "Industry 4.0". *Proc. Manuf.* 3, 417–424. doi: 10.1016/j.promfg.2015.07.200

Beke, E., Horvath, R., and Takacs-Gyorgy, K. (2020). Industry 4.0 and current competencies. *Naše Gospodarstvo Econ.* 66, 63–70. doi: 10.2478/ngoe-2020-0024

Bogoviz, A. V., Osipov, V. S., Chistyakova, M. K., and Borisov, M. Y. (2019). "Comparative analysis of formation of industry 4.0 in developed and developing countries," in *Proceedings of the industry 4.0: industrial revolution of the 21st century*, 155–164. doi: 10.1007/978-3-319-94310-7_15

Costan, E., Gonzales, G., Gonzales, R., Enriquez, L., Costan, F., Suladay, D., et al. (2021). Education 4.0 in developing economies: A systematic literature review of implementation barriers and future research agenda. *Sustainability* 13:12763. doi: 10.3390/su132212763

Da Costa, M. B., Dos Santos, L. M. A. L., Schaefer, J. L., Baierle, I. C., and Nara, E. O. B. (2019). Industry 4.0 technologies basic network identification. *Scientometrics* 121, 977–994. doi: 10.1007/s11192-019-03216-7

Das, S., Kleinke, D. K., and Pistrui, D. (2020). "Reimagining engineering education: Does industry 4.0 need education 4.0?," in *Proceedings of the 2020 ASEE virtual annual conference content access*. doi: 10.18260/1-2-35136

Ejsmont, K., Gladysz, B., and Kluczek, A. (2020). Impact of industry 4.0 on sustainability—bibliometric literature review. *Sustainability* 12:5650.

Fisk, P. (2017). *Education 4.0 the future of learning will be dramatically different, in school and throughout life*. Available online at: https://www.peterfisk.com/2017/01/future-education-young-everyone-taught-together/ (accessed July 25, 2022)

Gajdzik, B., Grabowska, S., Saniuk, S., and Wieczorek, T. (2020). Sustainable development and Industry 4.0: A bibliometric analysis identifying key scientific problems of the sustainable Industry 4.0. *Energies* 13:4254.

Galatro, D., Chakraborty, S., Yan, N., Goodarzi, N., Castrucci, J., and Saban, M. (2022). "Education 4.0: Integrating codes, standards, and regulations in the chemical engineering curriculum," in *Proceedings of the canadian engineering education association (CEEA)* (Toronto, Canada).

González-Pérez, L. I., and Ramírez-Montoya, M. S. (2022). Components of education 4.0 in 21st century skills frameworks: Systematic review. *Sustainability* 14:1493.

Grzybowska, K., and Łupicka, A. (2017). Key competencies for industry 4.0. *Econ. Manag. Innov.* 1, 250–253. doi: 10.26480/icemi.01.2017.250.253

Hecklau, F., Galeitzke, M., Flachs, S., and Kohl, H. (2016). Holistic approach for human resource management in Industry 4.0. *Proc. Cirp* 54, 1–6. doi: 10.1016/j.procir.2016.05.102

Hernandez-de-Menendez, M., Morales-Menendez, R., Escobar, C. A., and McGovern, M. (2020). Competencies for industry 4.0. *Int. J. Interact. Design Manuf.* 14, 1511–1524. doi: 10.1007/s12008-020-00716-2

Himmetoglu, B., Aydug, D., and Bayrak, C. (2020).). Education 4.0: Defining the teacher, the student, and the school manager aspects of the revolution. *Turk. Online J. Distance Educ.* 21, 12–28. doi: 10.17718/tojde.770896

Hirschi, A. (2018). The fourth industrial revolution: Issues and implications for career research and practice. *Career Dev. Q.* 66, 192–204. doi: 10.1002/cdq.12142

ITU (2022). *The ICT development index*. Available online at: https://www.itu.int/en/ITU-D/Statistics/Pages/IDI/default.aspx (accessed July 25, 2022)

Janik, A., and Ryszko, A. (2018). "Mapping the field of Industry 4.0 based on bibliometric analysis," in *Proceedings of the international business information management association conference (IBIMA)*, Vol. 32. doi: 10.3390/foods11193137

Jerman, A., Pejiæ Bach, M., and Aleksiæ, A. (2020). Transformation towards smart factory system: Examining new job profiles and competencies. *Syst. Res. Behav. Sci.* 37, 388–402. doi: 10.1002/sres.2657

Kipper, L. M., Iepsen, S., Dal Forno, A. J., Frozza, R., Furstenau, L., Agnes, J., et al. (2021). Scientific mapping to identify competencies required by industry 4.0. *Technol. Soc.* 64:101454. doi: 10.1016/j.techsoc.2020.101454

Lin, W. D., and Low, M. Y. H. (2021). "An integrated engineering education alignment model towards Industry 4.0," in *Proceedings of the international conference on industrial engineering and operations management*, 1204–1213.

Maisiri, W., and van Dyk, L. (2020). "Industry 4.0 competence maturity model design requirements: A systematic mapping review," in *In Proceedings of the 2020 IFEES world engineering education forum-global engineering deans council (WEEF-GEDC)* (Piscataway, NJ: IEEE), 1–6. doi: 10.1109/weef-gedc49885.2020.9293654

Mian, S. H., Salah, B., Ameen, W., Moiduddin, K., and Alkhalefah, H. (2020). Adapting universities for sustainability education in industry 4.0: Channel of challenges and opportunities. *Sustainability* 12:6100. doi: 10.3390/su12156100

Miranda, J., Navarrete, C., Noguez, J., Molina-Espinosa, J. M., Ramírez-Montoya, M. S., Navarro-Tuch, S. A., et al. (2021). The core components of education 4.0 in higher education: Three case studies in engineering education. *Comput. Electr. Eng.* 93:107278. doi: 10.1016/j.compeleceng.2021.107278

Neaga, I. (2019). "Applying Industry 4.0 and Education 4.0 to engineering education," in *Proceedings of the canadian engineering education association (CEEA)*. doi: 10.24908/pceea.vi0.13859

Oleksiyenko, A. (2021). "Global higher education and COVID-19: A virtual autoethnography of a faculty," in *Proceedings of the teaching and learning in higher education during COVID-19* (Milton Park: Routledge), 167–180. doi: 10.4324/9781003125921-16

Perry, B. L., Aronson, B., and Pescosolido, B. A. (2021). Pandemic precarity: COVID-19 is exposing and exacerbating inequalities in the American heartland. *Proc. Natl. Acad. Sci. U.S.A.* 118:e2020685118. doi: 10.1073/pnas.2020685118

Prifti, L., Knigge, M., Kienegger, H., and Krcmar, H. (2017). "A competency model for "industrie 4.0" employees," in *Proceedings der 13th internationalen tagung wirtschaftsinformatik (WI 2017)*, J. M. Leimeister, and W. Brenner (Hrsg.) eds (St. Gallen: University of St. Gallen), 46–60.

Ramirez-Mendoza, R. A., Morales-Menendez, R., Iqbal, H., and Parra-Saldivar, R. (2018). "Engineering Education 4.0:—proposal for a new Curricula," in *Proceedings of the 2018 IEEE global engineering education conference (EDUCON)* (Piscataway, NJ: IEEE), 1273–1282.

Ramírez-Montoya, M. S., Castillo-Martínez, I. M., Sanabria-Z, J., and Miranda, J. (2022). Complex thinking in the framework of Education 4.0 and Open Innovation—A systematic literature review. *J. Open Innov.* 8:4. doi: 10.3390/joitmc8010004

Reimers, F. M., and Marmolejo, F. J. (2022). *University and school collaborations during a pandemic: sustaining educational opportunity and reinventing education*. Berlin: Springer Nature, 357.

Silva, D. E., Lopes, T., Sobrinho, M. C., and Valentim, N. M. C. (2021). Investigating initiatives to promote the advancement of education 4.0: A systematic mapping study. *CSEDU* 1, 458–466. doi: 10.5220/0010439704580466

Uhlemann, J., Costa, R., and Charpentier, J. C. (2019). Product design and engineering in chemical engineering: Past, present state, and future. *Chem. Eng. Technol.* 42, 2258–2274. doi: 10.1002/ceat.201900236

United Nations (2022a). *Human development index. Human development reports*. Available online at: https://hdr.undp.org/data-center/human-development-index#/indicies/HDI (accessed July 25, 2022).

United Nations (2022b). *The 17 goals | sustainable development*. New York, NY: United Nations.

Van Eck, N. J., and Waltman, L. (2007). "VOS: A new method for visualizing similarities between objects," in *Proceedings of the 30 th annual conference of the gesellschaft für klassifikation eV, freie universität Berlin: Advances in Data Analysis* (Berlin: Springer Berlin Heidelberg), 299–306.

Verner, J. M., Brereton, O. P., Kitchenham, B. A., Turner, M., and Niazi, M. (2012). "Systematic literature reviews in global software development: a tertiary study," in *Proceedings of the 16th international conference on evaluation & assessment in software engineering (EASE 2012)* (London: IET), 2–11. doi: 10.1049/ic.2012.0001

VOSviewer (2022). *Visualizing scientific landscapes*. Available online at: https://www.vosviewer.com/ (accessed July 25, 2022)

Waaijer, C. J., van Bochove, C. A., and van Eck, N. J. (2011). On the map: Nature and Science editorials. *Scientometrics* 86, 99–112.

Wakamo, B. (2022). *Get the facts on inequality and covid-19*. Available online at: https://inequality.org/facts/inequality-and-covid-19/ (accessed July 25, 2022).

Wallner-Drewitz, T., and Wagner, G. (2016). "Academic Education 4.0," in *Proceedings on the international conference on education and new developments Ljubljana 2016*.

Wilke, A., and Magenheim, J. (2017). "Requirements analysis for the design of workplace-integrated learning scenarios with mobile devices: Mapping the territory for learning in industry 4.0," in *Proceedings of the 2017 ieee global engineering education conference (EDUCON)* (Piscataway, NJ: IEEE), 476–485. doi: 10.1109/educon.2017.7942890

Frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 31 May 2023
DOI 10.3389/feduc.2023.1161772

Check for updates

OPEN ACCESS

EDITED BY
Wang-Kin Chiu,
The Hong Kong Polytechnic University, China

REVIEWED BY
Joseph R. Boyle,
Temple University, United States
Kit W. Cho,
University of Houston–Downtown,
United States

*CORRESPONDENCE
Carla M. Firetto
✉ cfiretto@asu.edu

RECEIVED 08 February 2023
ACCEPTED 21 April 2023
PUBLISHED 31 May 2023

CITATION
Firetto CM, Starrett E, Montalbano AC, Yan L, Penkrot TA, Kingsbury JS and Hyatt J-PK (2023) The impact of effective study strategy use in an introductory anatomy and physiology class. *Front. Educ.* 8:1161772.
doi: 10.3389/feduc.2023.1161772

The impact of effective study strategy use in an introductory anatomy and physiology class

Carla M. Firetto[1]*, Emily Starrett[1], Amy Collins Montalbano[2], Lin Yan[1], Tonya A. Penkrot[3], Jeffrey S. Kingsbury[3] and Jon-Philippe K. Hyatt[3]

[1]Mary Lou Fulton Teachers College, Arizona State University, Tempe, AZ, United States, [2]Mathematics Department, Northwest Vista College, San Antonio, TX, United States, [3]College of Integrative Sciences and Arts, Arizona State University, Tempe, AZ, United States

Introductory courses in biology often act as a gateway for students seeking careers in healthcare and science-related fields. As such, they provide a prime entry point for innovations seeking to enhance students' learning of foundational content. Extant innovations and interventions have been found to positively impact students' study strategy use with concomitant impacts on course exams and grades. These innovations, however, often have associated time and other costs, which may ultimately limit more widespread use. Our study builds on prior findings by exploring the extent to which students evidence increased use of effective study strategies after engaging in a brief (i.e., 15-min), online module requiring no financial cost for students or time commitment from instructors, and whether changes in students' use of effective study strategies are associated with changes in exam performance. The present study employed a brief, online module designed to support undergraduate students' (n=98) use of effective study strategies in an introductory human anatomy and physiology course. Through a pretest-posttest design, students described the strategies they used to study and completed four cognitive and metacognitive subscales before and after engaging in a brief, online module designed to teach them about effective study strategies. Results were somewhat mixed: students evidenced a modest, statistically significant increase in the number of strategies used and changes in strategy use were associated with changes in exam score only for some measures. Notably, this relationship was not moderated by GPA, suggesting that the strength of the relationship between changes in strategy use and changes in exam scores were not different depending on students' levels of prior academic performance. Taken together, the innovation was associated with increases in students' exam scores, irrespective of GPA, but future research should explore the refinement and extension of the innovation to explore ways that increase efficacy and impact while still balancing sustainable implementation to account for challenges associated with instructor supervision and training, financial costs, and students' time.

KEYWORDS

self-regulated learning, study strategies, online intervention, undergraduate students, learning

Introduction

There is often a disconnect between the effective study strategies and learning techniques that educational experts know to be most effective for learners (Kornell and Bjork, 2007; Karpicke et al., 2009; Weinstein et al., 2010; Hartwig and Dunlosky, 2012; Dunlosky et al., 2013; Vemu et al., 2022) and the ones that students employ when they study. Practice testing and distributed practice, for example, are more efficacious for learning and comprehension than summarizing, highlighting, or rereading (Dunlosky et al., 2013; Adesope et al., 2017). Despite this evidence, undergraduate students often rely on less effective techniques, such as rereading (Karpicke et al., 2009) and waiting to study until right before a test (i.e., massed practice or "cramming," Blasiman et al., 2017). Ultimately, many students enter colleges and universities underprepared for how to learn (Kiewra, 2002; Wingate, 2007; Kritzinger et al., 2018). Recognizing this disconnect, educational researchers have long endeavored to support students' use of effective study strategies (Hattie et al., 1996) and delineate ways to help college students use these strategies successfully (Cook et al., 2013; Zhao et al., 2014; Broadbent and Poon, 2015; Muteti et al., 2021; Theobald, 2021).

Undergraduate students enrolled in introductory biology courses are a particularly important subpopulation with contextual demands that set them apart from other majors within the university and may particularly benefit from engaging in more effective study strategies (Roediger and Butler, 2011; Hartwig and Dunlosky, 2012; Blasiman et al., 2017; Kritzinger et al., 2018; Vemu et al., 2022). For example, students taking biology-related courses, such as human anatomy and physiology, are often seeking health science careers, and thus, these courses may serve as a "gateway" into those careers, in effect, granting or limiting access (Koch, 2017; Hensley et al., 2021; Muteti et al., 2021). Students that employ better study strategies may be more likely to excel in the course, learn more, and ultimately be retained in their chosen field. For example, Schneider and Preckel (2017) found in their meta-analysis that students utilizing study strategies, such as elaboration or retrieval practice, performed at higher rates than peers who do not use those strategies. This notion is further impacted by the finding from a study by Marbach-Ad et al. (2016) where biology students that had lower grade point averages (GPAs) placed a higher value on retention skills (e.g., rote memorization of concepts, such as listing the bones in the body) over transfer skills (e.g., deeper understanding of content, such as application of concepts to health care contexts). Further, Ley and Young (1998) found that students who entered college underprepared (i.e., classified as taking a developmental or remedial class) not only used fewer total strategies but also used them with less consistency than their regular admission peers. Given that students with the greatest need may use strategies less frequently and have a greater predisposition toward reliance on less effective rehearsal strategies, they may benefit from targeted support.

A growing body of innovations designed to support biology students' study strategy use has emerged over the past 2 decades (Minchella et al., 2002; Sebesta and Bray Speth, 2017; Bernacki et al., 2020; Hensley et al., 2021). Our study builds on prior findings by exploring the extent to which students evidence increased use of effective study strategies after engaging in a brief (i.e., 15-min), online module requiring no financial cost for students or time commitment from instructors, and whether students' increased use of effective study strategies is associated with increased exam performance. We also examine whether the strength of that potential relationship is moderated by GPA.

Approaches to supporting students' study strategy use have emerged from different frameworks, including self-regulated learning (SRL; Zimmerman, 2000, 2002) and desirable difficulties (Bjork and Bjork, 2011). While SRL has been conceptualized in multiple ways, many researchers have centered the three-phase approach forwarded by Zimmerman (2000, 2002), whereby learning occurs through a cyclical process involving forethought (i.e., *before*), performance (i.e., *during*), and self-reflection (i.e., *after*). Each of these interdependent phases plays a critical role in how individuals learn. For example, students' self-motivation beliefs are an important subprocess of the forethought phase. The extent to which students come to an intervention already possessing intrinsic interest or value, self-efficacy, or mastery-oriented goals will impact their engagement in sustained study efforts and learning (Eccles, 1983; Bandura, 1986; Ames and Archer, 1988; Zimmerman and Schunk, 1989; Zimmerman, 2002).

Many students enrolled in introductory biology courses are pursuing careers in healthcare and science-related careers—careers that ultimately require mastery of such content (Koch, 2017; Hensley et al., 2021; Muteti et al., 2021). Therefore, students may recognize the value of learning about anatomy and physiology (Sullins et al., 1995), positively impacting the forethought phase. Yet students often cannot identify (i.e., *forethought*), deploy (i.e., *performance*), or evaluate (i.e., *self-reflection*) strategies that they do not know, and thus, SRL is constrained by the repertoire of study strategies and progress monitoring approaches that learners possess (Zimmerman and Schunk, 1989; Bernacki et al., 2020). This is particularly critical for introductory biology students, as Sebesta and Bray Speth (2017) found that they not only have limited knowledge of SRL strategies, but they may also be unable to properly implement them. Their findings also revealed a link between SRL strategy use and achievement that was previously found in other science content areas (see Lopez et al., 2013).

Not all study strategies are equally effective. Bjork and Bjork (2011) argue that strategies that induce *desirable difficulties* yield greater cognitive understanding and better enable encoding and retrieval processes. The desirable difficulties framework asserts that employing more effortful and active strategies (e.g., interleaving, spaced studying, using quizzes, or practice tests to study material) cultivates longer and deeper comprehension (Bjork and Bjork, 2011). Walck-Shannon et al. (2021) leveraged the desirable difficulties framework to examine the relationship between study strategies and performance on exams for introductory biology students. They found that students who used a greater number of active study strategies (e.g., explaining concepts, self-quizzing, and drawing diagrams) scored higher than students who used fewer active strategies or passive strategies (i.e., read textbooks, rewrote notes, and watched lectures). Each additional active strategy that students used was associated with an increase of about 2–3% on the respective exams. Further, Kritzinger et al. (2018) found that ability and willingness to persist through challenges were more evident in higher-performing students and can be predictive of student success, underscoring the impact of prior performance.

The use of interventions to support introductory biology students is not new. Minchella et al. (2002) investigated the impact of a semester-long, one-credit, biology seminar designed to help first-year students transition to college and increase their academic success. Academic advisors and a team of undergraduate teaching interns assisted first-year students through problem-solving sessions (e.g., class time devoted to modeling and teaching problem-solving strategies for the concurrent biology lab), as well as discussions and lectures, with an emphasis on fostering collaborative peer support. Throughout the semester, students developed time management systems, learned strategies to help them succeed in biology, and had time to visit research laboratories. Overall, these activities helped students build realistic expectations of a career in the field of biology. The seminar course resulted in positive outcomes including increased grades, student satisfaction, and retention in the department.

More recently, Bernacki et al. (2020) implemented a 2-h, self-guided, online training course embedded within a biology seminar. The goal of this study was to examine if a "Learning to Learn" course could change undergraduate biology students' study habits and improve their coursework performance. The intervention contained three modules designed to teach and model the effectiveness of different learning strategies by providing opportunities for students to read and practice not only using the strategies but adapting them to their needs. The modules ended with identifying resources provided within the biology seminar's learning management system (LMS) to help future learning. The modules had a statistically significant impact on student behavior (e.g., students utilized more self-assessment, planning, and self-monitoring resources than students that did not participate in the modules, as measured by monitoring the LMS traffic) and academic performance (e.g., students scored higher on exam scores than those that did not participate).

Interventions have the potential to yield increased learning outcomes for undergraduate students in biology-related courses and beyond. However, many of these approaches require a significant financial and time investment (e.g., training of instructors, days or hours required for students to complete the module). Comparatively fewer approaches have emphasized more sustainable implementation (e.g., brief, online, and low resource). One notable recent exception centered on a single, brief (e.g., 15 min) instructor-created presentation and discussion that focused on three high-impact strategies (Vemu et al., 2022). In their intervention, Vemu et al. (2022) encouraged students to engage in high-impact, effective study strategies (i.e., spacing, self-testing, and drawings or models) at the beginning of the semester. While there was no statistically significant growth in students' use of key strategies from the beginning to the end of the semester, students that reported using spacing and drawing strategies by the end of the semester had higher grades.

Our study advances extant research by exploring the extent to which a study strategy intervention that is not only brief (i.e., 15-min) but also instructor-independent (e.g., not requiring additional instructor/course time) can yield a positive impact on student learning outcomes for students in an introductory anatomy and physiology course.

The present study

We employed a one-group, repeated measures design, such that all students engaged in the brief, online module between exam 2 and exam 3. This design allowed us to look at students' strategy use over time, as well as the extent to which strategy use was linked to exam score. Further, it enabled us to look at whether that potential relationship was moderated by students' prior academic performance, such that students with different levels of prior academic performance (e.g., comparatively higher or lower GPA) have a stronger or weaker relationship between the changes in their strategy use and exam score differences.

> RQ1: Do students evidence greater use of effective study strategies after participating in a brief, online module, as evidenced by descriptions of their strategy use and ratings on cognitive and metacognitive strategies subscales?
>
> RQ2: (a) Are changes in students' use of effective study strategies associated with changes in exam score and (b) is this relationship moderated by self-reported GPA, as evidenced by descriptions of their strategy use and ratings on cognitive and metacognitive strategies subscales?

Materials and methods

Participants, context, and design

Undergraduate students were recruited from three, large sections of an introductory human anatomy and physiology course taught in the spring semester by two instructors at a large public, Hispanic-serving University in the southwestern United States. The course covered aspects related to the structure and function of the human body, including cells and tissues as well as the integumentary, skeletal, muscular, and nervous systems. All three sections of the course were taught predominantly via traditional lecture with an associated lab component. Participating students ($n = 98$) made up about 16% of the total number of initially enrolled students across the three sections (i.e., between 140 and 240 students per section, not accounting for those who withdrew from the course).

Participants (women, $n = 74$; men, $n = 22$; nonbinary, $n = 1$; did not respond, $n = 1$) were mostly (85.6%) between 18 and 21 years old ($M = 19.95$, $SD = 2.16$). Students identified as White[1] (55.0%), Hispanic (28.6%), Asian (14.3%), Black (9.2%), American Indian (3.1%), Pacific Islander (2.0%), or elected not to report their race (2.0%). Over half were students in their first year of college (57.1%) with the remaining participants in their second (35.7%) or third (7.2%) year. Almost all of the participants (90%) expressed that they were taking the course at least partly because it was a required course for their major, but a substantial portion also noted that they were interested in learning the course content (39%) or that it would help them with their future career (61%).

Human subjects approval was obtained prior to conducting the study (#STUDY00008599), and all participants consented to participate in the research before beginning the first survey. APA ethical standards were followed throughout the duration of the research. Students were offered 2% extra credit in their course as compensation for completing the study. All but one participant granted permission to include exam grades as part of our data, thus that individual was excluded from analyses that involved exam grades.

1 Total does not equal 100%, as students were permitted to select multiple race identifiers.

We intentionally employed a one-group, pretest-posttest design that invited all students enrolled in the class to engage in the module midway through the semester. This timing allowed for a more stable measure of students' typical strategy use at pretest, having already experienced one exam before reporting strategy use on the second exam (see also Bernacki et al., 2020). Sebesta and Bray Speth (2017) referred to this as the "settling in" (p. 9) of strategy use occurring after the second exam. Further, given the emphasis on a brief intervention, we were particularly interested in examining the impact on students' study strategy use immediately after the module (i.e., the exam that followed several weeks later), where it would most likely be detected, before examining the potential for delayed impact (i.e., the final exam).

Materials

Brief, online module

The brief, online module focuses on six study strategies that have been largely established in the literature as effective but not commonly discussed in classrooms (Pomerance et al., 2016) or used by students (Dunlosky et al., 2013; Weinstein et al., 2023a): spaced practice (Benjamin and Tullis, 2010), retrieval practice (Roediger et al., 2011), elaboration (McDaniel and Donnelly, 1996), interleaving (Rohrer, 2012), concrete examples (Rawson et al., 2014), and dual coding (Mayer and Anderson, 1992). In alignment with our theoretical framing, Dunlosky et al. (2013) identified these strategies among those that can help students improve their comprehension and application of concepts, allowing individuals to better engage in SRL (e.g., use more effective strategies in the performance phase; Zimmerman, 2000). Likewise, Bjork and Bjork (2011) noted several of these as active study strategies that elicit desirable difficulties. Further, using these strategies in combination can help solidify the study process, given their complementary nature. For example, spaced practice focuses on spreading out study sessions, whereas dual coding and concrete examples emphasize how one can effectively study during those spaced study sessions (Weinstein et al., 2023a). Similarly, retrieval practice can not only help improve the ability to recall information, but also when spaced out over time, it can aid transfer of knowledge to new contexts (Butler, 2010).

All students participated in a brief (i.e., approximately 15 min), two-part module where they (a) learned about the six study strategies and (b) reflected on how they could use two of the strategies in their human anatomy and physiology class. First, students watched a video (8.5 min; Memorize Academy [Username] in collaboration with the Learning Scientists, 2016) that overviewed all six strategies. The video was produced in collaboration with The Learning Scientists,[2] cognitive psychologists that study the science of learning, and addressed both how to use each strategy as well as an overview of research that supports their benefits on learning. Students were unable to proceed to the next page of the survey until the duration of the video had elapsed. Then, students ranked the strategies based on what they were most interested in learning about in more depth. For their two highest interest strategies, students spent 3–5 min reviewing the associated infographic (Weinstein et al., 2023b) and writing a detailed plan for how they could use that strategy to study for their human anatomy and physiology class (Figure 1). While students' detailed plans were not evaluated as part of the data, the authors verified that students responded to the planning prompt.

Quantity of effective study strategies used

After both exam 2 and exam 3, students responded to a series of open-ended questions (e.g., "please describe all of the strategies you used to study in as much detail as possible") asking them to describe how they studied for the exam they just took (Figure 2). The responses were coded based on whether students described using each of the six different study strategies across their responses (i.e., used = 1, not = 0). A quantity score was also calculated for each student based on the total number of effective study strategies they described using for exam 2 (i.e., before the module) and for exam 3 (i.e., after the module). Scores could range from 0 (i.e., no effective strategies) to 6 (i.e., all effective strategies). For example, one student described their studying by noting, "I used quizlet to memorize terms, flash cards to test myself[,] and I drew myself pictures of types of tissues, bones, and diagrams[,] such as [a] hair follicle[,] we needed to know to help myself study and understand the structures." This response represents a score of 2, as the student described using both retrieval practice (i.e., quizlet and/or flash cards) and dual coding (i.e., drawing pictures and/or diagrams). All responses were coded by the third author and 20% of the responses were then checked by the first author for fidelity to the scoring rubric and interrater consistency. Interrater agreement was checked separately for the identification of each strategy within a student's response. This process allowed us to ensure that agreement was sufficient for each strategy independently [i.e., ICC (2), absolute agreement, single measure >0.698], as well as for overall quantity score [i.e., the total sum of all effective strategies used; ICC (2), consistency, single measure = 0.888]. We also calculated a strategy use change score (i.e., the quantity of strategies students described using at exam 3 minus the quantity of strategies students described using at exam 2) to gauge the extent to which students' use of effective study strategies changed over time.

Cognitive and metacognitive strategies subscales

Students completed the Motivated Strategies for Learning Questionnaire (Pintrich et al., 1991), which included the cognitive and metacognitive strategies (CAMS) subscales, after both exam 2 and exam 3. Participants responded to each of the statements (e.g., "When reading for this course, I make up questions to help focus my reading.") on a 7 (i.e., *very true of me*) to 1 (i.e., *not at all true of me*) Likert-type scale, and scores for each subscale were calculated averaging across all items associated with the respective subscale. Given the focus of the module, we report only data pertaining to four of the CAMS subscales (elaboration, $\alpha_{pre} = 0.665$; $\alpha_{post} = 0.702$; organization, $\alpha_{pre} = 0.385$; $\alpha_{post} = 0.567$; critical thinking, $\alpha_{pre} = 0.659$; $\alpha_{post} = 0.500$; and metacognitive self-regulation, $\alpha_{pre} = 0.733$; $\alpha_{post} = 0.834$). Notably, Cronbach alpha values for three of the subscales (i.e., elaboration, organization, and critical thinking) were right at or below the threshold of $\alpha > 0.7$, potentially due to the low number of items combined with the somewhat modest sample size. For each subscale, correlations between the two administrations (i.e., at exam 2 and at exam 3) were all statistically significant and positive (all $rs > 0.415$), providing additional evidence of test–retest reliability of the subscale scores.

2 www.learningscientists.org

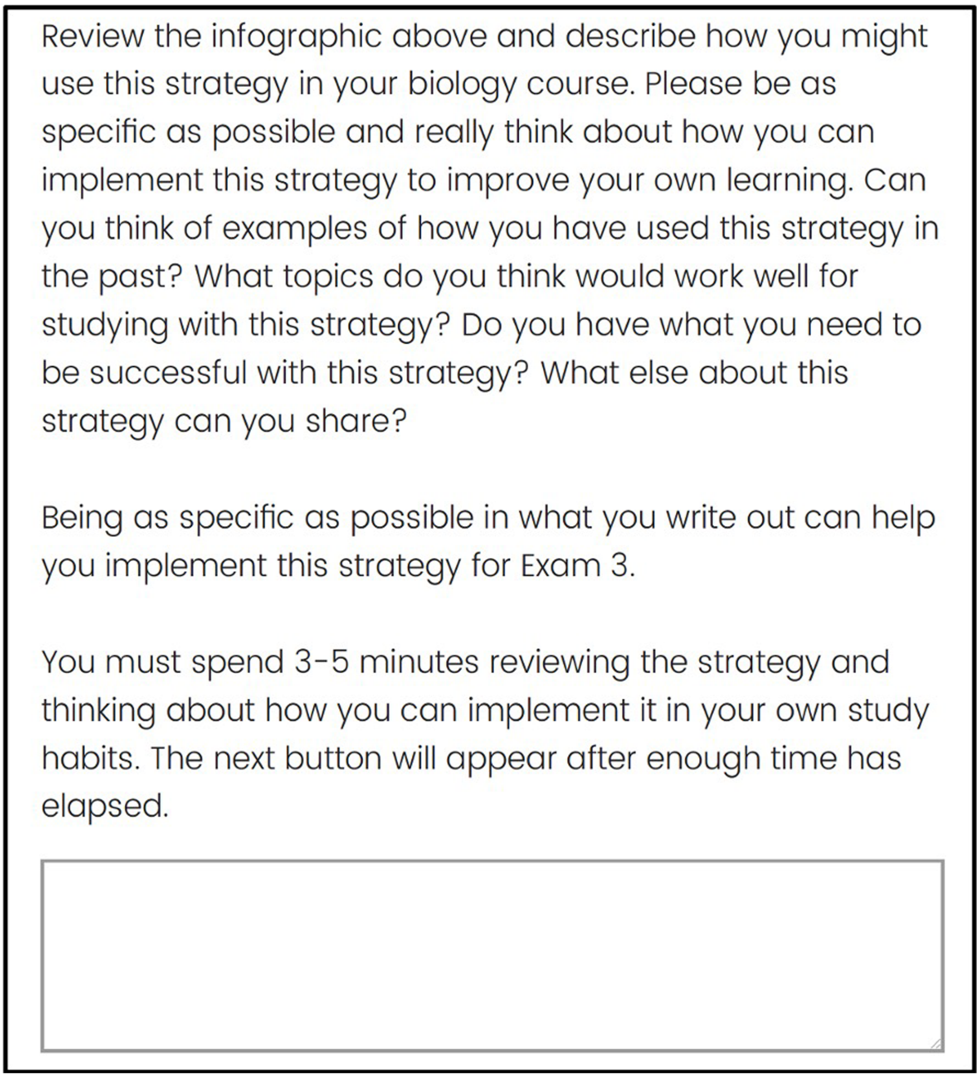

FIGURE 1
Strategy implementation prompt example from module. Infographics referred to in the prompt were produced by Weinstein et al., 2023b and are available at: https://www.learningscientists.org/downloadable-materials.

Demographic information and self-reported GPA

At the end of the second survey, participants completed a brief demographic questionnaire (e.g., age, gender, race, and enrollment), and participants were asked to self-report their college GPA. Additionally, several questions also focused on participants' motivations for taking the course (e.g., their plans after graduation, whether the course was required for their program or major).

Exam scores

Four exams were administered in the course, roughly 4 weeks apart. Each exam was worth 40 points and together they contributed to 50% of students' total course grade. The content assessed in each exam was independent and non-cumulative, that is, exams targeted only the content learned over the preceding 4 weeks. Difficulty was not equated between the four exams. Both instructors reported overall average scores for the four exams, indicating a progressive increase in difficulty over time (i.e., each exam had a lower average percentage than the preceding one). Specifically, the drop from exam 2 to exam 3 was 2% for one instructor and 3% for the other. Instructors noted that students typically could draw more from prior knowledge based on content learned earlier in the course (e.g., prior college chemistry course or advanced biology course in high school) than later on in the semester. Scores for the first, second, and third course exams were obtained from course instructors for students that consented to allow grades to be used as part of the research (i.e., all but one). To address RQ2, we also calculated an exam change score (i.e., exam 3 minus exam 2). This allowed us to look specifically at the extent to which changes in strategy use were associated with increased or decreased performance on the exam.

Procedures

Students were invited to participate in the research immediately after receiving their grades for exam 2. After consenting to participate in the research, students completed the pretest survey (see Figure 3),

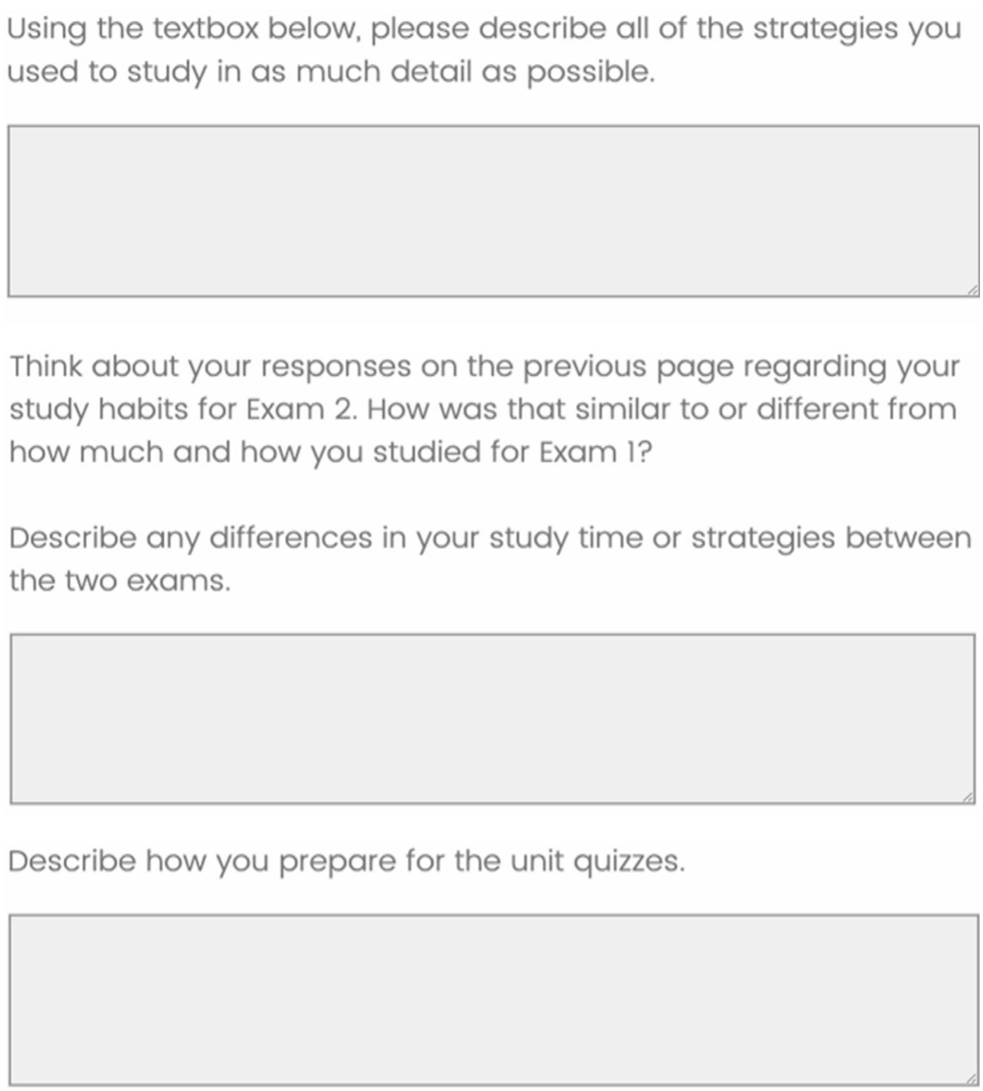

FIGURE 2
Prompts to gather quantity of effective study strategies used.

which included (a) their descriptions of how they studied for the exam, (b) the CAMS subscales, and (c) the brief, online module. Two weeks later, an email went out to all students who completed the pretest survey that included a reminder link to the video with a note prompting them to use the study strategies while preparing for exam 3. Students were emailed the link for the posttest survey immediately after the grades for exam 3 were posted. The second survey included the same measures as the first (i.e., *a* and *b* above), and it also included a series of demographic and motivation questions, including a self-report of their current GPA.

Statistical analysis

Given the ordinal nature of the quantity of effective study strategies described using (i.e., 0–6), we used Related-Samples Wilcoxon Signed Rank test to examine RQ1 and the changes in study use from exam 2 to exam 3. Prior to analyzing the data, we examined the distribution of the differences, which revealed a symmetrically shaped distribution, thus meeting the requisite assumption for interpreting the results of this test. In contrast, we used paired-samples *t*-tests to determine whether the mean differences on the four CAMS subscales from exam 2 to exam 3 were statistically significant, given the continuous nature of the subscale scores. After examining the boxplots for each respective subscale analysis, outliers that were more than 1.5 box-lengths from the edge of the box were removed, ranging from no outliers on the organization subscale to five outliers on the elaboration subscale. The assumption of normality was not violated for any of the subscales (i.e., Shapiro–Wilk's test, all $ps > 0.186$). For RQ2, we employed PROCESS v4.1 macro of Hayes (2021) in SPSS to gauge whether a change in effective strategies (i.e., measured by the quantity of strategies participants described using and the four CAMS subscales) was associated with a change in exam performance (i.e., a higher grade on exam 3 than exam 2) and whether the strength of that relationship was moderated by students' self-reported GPA.

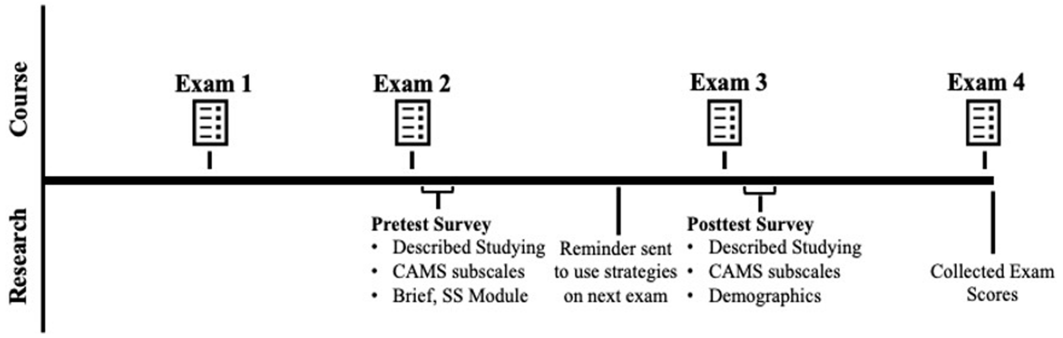

FIGURE 3
Timeline of procedures.

TABLE 1 Descriptive statistics related to key variables.

	For exam 2 (Pre-module survey)				For exam 3 (Post-module survey)				Changes (Exam 3—Exam 2)			
	Min	Max	Mean	*SD*	Min	Max	Mean	*SD*	Min	Max	Mean	*SD*
Quantity of Study Strategy (*n* = 98)	0.00	3.00	0.82	0.75	0.00	5.00	1.09	1.01	−3.00	3.00	0.28	0.99
CAMS study strategies subscales (*n* = 98)												
Elaboration (items = 6)	2.00	7.00	4.49	1.18	1.17	7.00	4.68	1.21	−3.00	3.83	0.19	1.29
Organization (items = 4)	2.50	7.00	4.80	1.11	1.75	7.00	4.85	1.21	−2.50	2.75	0.05	1.03
CT (items = 5)	1.20	6.80	3.55	1.07	1.00	7.00	3.77	1.03	−3.20	3.00	0.24	1.06
Meta SR (items = 12)	1.25	6.83	4.51	0.91	1.25	7.00	4.65	1.00	−1.92	2.50	0.13	0.72
Exam scores (2, *n* = 96; 3, *n* = 95)	12.43	38.83	28.19	6.01	7.22	38.40	26.22	6.35	−14.33	18.97	−2.04	5.11
GPA (*n* = 94)	-	-	-	-	2.42	4.14	3.61	0.40	-	-	-	-

CAMS, cognitive and metacognitive strategies; CT, critical thinking; Meta SR, metacognitive self-regulation; and GPA, grade point average. GPA data are reported in the middle set of columns, as students reported GPAs at the end of the second survey after the module.

Results

Descriptive statistics

Students, on average, evidenced modest strategy use for both exam 2 and exam 3, as evidenced by the quantity of study strategies reported and CAMS study strategy subscale scores (Table 1). As mentioned previously, exam 3 had a higher difficulty for students than exam 2 (i.e., the section averages for exam 3 were between 2 and 3% lower than on exam 2 scores). However, participating students only scored about two points (i.e., 0.5% of the 40-point exam) lower on exam 3 than on exam 2. GPA was overall notably high; only five students reported a GPA below 3.0, while 25 reported a GPA at or above 4.00 (i.e., grades of A+ are weighted at 4.33). We address issues related to the overall high GPA scores and the decision to collect GPA via self-report in greater detail in the discussion.

Frequency counts for each of the six study strategies are noted in Table 2. Prior to the brief, online module, retrieval practice was noted most frequently as the strategy students described using. After the module, there was a very small increase (i.e., between 3 and 8) in the number of students who reported using each of the strategies, except for dual coding. Notably, students selected spaced practice and retrieval practice as the ones they were most interested in learning about—they were selected almost twice as frequently as the other strategies—these were also strategies among those that students most often reported using prior to watching the video (i.e., they described using them for exam 2).

There was a statistically significant, positive correlation between the quantity of study strategies students used for exam 2 and exam 2 scores ($r = 0.273$, $p = 0.007$). Of note, there was no statistically significant correlation between the quantity strategies used for exam 2 and the other exam scores (i.e., exam 3, $r = 0.142$, $p = 0.169$; exam 1, $r = 0.167$, $p = 0.105$). This pattern, however, did not hold for the quantity of study strategies students used for exam 3. There was no significant correlation with any of the exams, including exam 3, $r = 118$, $p = 0.255$, although the correlation between the quantity of study strategies used at exam 3 was higher for exam 3 than it was for exam 1, $r = 0.017$, $p = 0.870$, or exam 2, $r = -0.012$, $p = 0.906$. There

TABLE 2 Frequency counts for the six strategies.

Study strategy ($n=98$)	Described using for exam 2 (freq)	Described using for exam 3 (freq)	Highest interest strategies (first or second choice)
Spaced practice	11	19	57
Retrieval practice	48	53	46
Elaboration	8	16	26
Interleaving	1	4	21
Concrete examples	0	5	26
Dual coding	12	10	20

were no significant correlations between any of the CAMS subscales at exam 2 and any of the exams, and for exam 3, only the metacognitive self-regulation subscale had a statistically significant positive correlation with the associated exam. These correlations suggest a limited pattern whereby students' use of effective study strategies was associated with higher exam scores (see Table 3).

Changes in study strategy use

To gather a more comprehensive understanding regarding changes in students' study strategy use over time, we analyzed RQ1 by looking at two different indicators of strategy use. First, we examined changes in the quantity of effective study strategies based on students' descriptions. Of the 98 participating students, 41 described using a greater number of effective study strategies in preparation for exam 3 than they did in preparation for exam 2 (i.e., after engaging with the module), 37 described using the same number, and only 20 described using fewer effective study strategies. Altogether, students evidenced a statistically significant median increase in the number of strategies from exam 2 to exam 3, $z=2.75$, $p=0.006$. Notably, however, the median number of strategies students used was the same ($Mdn=1$) both before and after the module.

Additionally, we also looked at changes in students' responses to the associated CAMS subscales. For both the critical thinking, $t(95)=2.02$, $p=0.046$, Cohen's $d=0.207$, and metacognitive self-regulation, $t(94)=2.16$, $p=0.033$, Cohen's $d=0.223$, subscales, there was a statistically significant mean increase over time in line with a small effect. No differences were detected for either the elaboration or organization subscales (both $ps>0.196$, Cohen's $d<0.135$).

Impact of study strategy change

When looking at the impact of study strategy change, the results revealed that the overall model (i.e., change in study strategies that students described using predicting change in exam score and accounting for GPA) was statistically significant, $F(3,87)=2.92$, $p=0.0384$, $R^2=0.09$. As predicted, the change in the number of strategies used was associated with a statistically significant change in exam score, $b=1.34$, $t(87)=2.58$, $p=0.012$, revealing that every additional strategy used was associated with an increase of 1.34 exam points (i.e., out of 40 points total) for those scoring at the grand mean of GPA. GPA did not directly predict exam score change, $b=0.83$, $t(87)=0.61$, $p=0.544$, and there was no interaction between changes in strategy use and GPA, $b=1.96$, $t(87)=1.23$, $p=0.114$. As such, the association between strategy use and scores on the exam was consistent across students, irrespective of their GPA.

Additionally, we examined changes in students' study strategy use as evidenced by their scores on the four CAMS subscales. However, the results revealed that none of the overall models were statistically significant [e.g., metacognitive self-regulation, $F(3,87)=1.44$, $p=0.237$, $R^2=0.05$; critical thinking, $F(3,87)=1.59$, $p=.197$, $R^2=0.05$].

Discussion

Drawing from extant interventions and grounded in the literature of SRL (Zimmerman, 2000, 2002) and desirable difficulties (Bjork and Bjork, 2011), the present study centered around examining an innovation for students taking an introductory human anatomy and physiology course to potentially increase their effective strategy use. Our approach offered a unique contribution in that it was designed for sustainable use (i.e., took only 15 min of students' time to complete, was completed outside of class time and online via a link, required no extra materials or costs for students, and did not involve any instructor time).

We employed two different indicators of students' strategy use. First, by systematically coding students' descriptions of their studying, we were able to measure the degree to which students used effective strategies in a way that was sensitive to the six specific strategies embedded in the module. Second, by using the CAMS subscales of the Motivated Strategies for Learning Questionnaire, we also gathered complementary measures of strategy use via a well-established measurement tool (Pintrich et al., 1991; Kritzinger et al., 2018).

Ultimately, we found modest, statistically significant increases on some indicators of students' strategy use. On average, students described using more effective study strategies, as well as greater critical thinking and metacognitive self-regulation, but there were no differences detected for two of the subscales (i.e., elaboration and organization). Additionally, there was limited evidence about the association between changes in strategy use and changes in exam scores, and GPA did not moderate this relationship.

Need for briefer, sustainable innovations

Numerous interventions have been designed to successfully support undergraduate students' self-regulated learning and study strategy use (Hattie et al., 1996; Minchella et al., 2002; Roediger and Butler, 2011; Hartwig and Dunlosky, 2012; Cook et al., 2013; Zhao et al., 2014; Broadbent and Poon, 2015; Blasiman et al., 2017; Sebesta and Bray Speth, 2017; Bernacki et al., 2020; Hensley et al., 2021; Muteti et al., 2021; Theobald, 2021; Vemu et al., 2022). Each is comprised of a unique composition of features that make up how it is implemented and enacted, specifying both modality (e.g., online vs. face-to-face) and intensity (e.g., long vs. short). These features necessarily impact both the possibility for scalable implementation (e.g., the ability for

TABLE 3 Correlation matrix for key variables.

	1	2⁺	3⁺	4#	5#	6#	7#	8	9	10
1. GPA	-	-	-	0.139	0.034	−0.158	0.146	-	-	-
2. SS quant exam 2⁺	0.222*	-	-	0.002	0.173	0.002	0.124	-	-	-
3. SS quant exam 3⁺	0.094	0.325**	-	0.146	0.257*	−0.152	0.164	-	-	-
4. CAMS Elab#	−0.052	0.142	0.164	**0.415****	0.489**	0.330**	0.496**	−0.010	0.048	0.008
5. CAMS Org#	−0.027	0.157	0.319**	0.571**	**0.607****	0.201*	0.518**	−0.008	0.070	0.022
6. CAMS CT#	−0.082	0.140	−0.011	0.490**	0.408**	**0.491****	361**	−0.072	−0.061	−0.089
7. CAMS Meta SR#	0.115	0.226*	0.291**	0.630**	0.619**	0.567**	**0.723****	0.082	0.117	0.182
8. Exam score 1	0.452**	0.167	0.017	−0.079	−0.122	−0.021	0.113	-	-	-
9. Exam score 2	0.526**	0.273**	−0.012	−0.086	−0.132	−0.091	0.092	0.659**	-	-
10. Exam score 3	0.545**	0.142	0.118	−0.020	−0.039	0.000	0.277**	0.627**	0.660**	-

⁺Spearman's Rho Correlations are reported for pairs that included "quantity" variables; all others reported are Pearson Correlations.
*$p < 0.05$; **$p < 0.001$ #Correlations with and between CAMS subscale scores for exam 2 are reported above the diagonal and correlations with and between CAMS subscale scores for exam 3 are reported below the diagonal. Values on the diagonal, noted in boldface type, represent correlations for each subscale between administration times (i.e., strategy subscale at exam 2 with that same strategy subscale at exam 3). SS quant, study strategy quantity; Elab, elaboration; Org, organization; CT, critical thinking; and Meta SR, metacognitive self-regulation.

other instructors to implement the intervention), as well as the likelihood for generalizability of results (e.g., whether the benefits to students are believed to apply in other contexts).

While longer interventions have shown greater efficacy (Dignath and Büettner, 2008), there are inherent challenges that come with them. Undergraduate students are faced with increasing competing demands for their time (e.g., balancing school and work) that may preclude their participation in longer-duration interventions. Shorter interventions can still be effective. The digital skills intervention designed by Bernacki et al. (2020), for example, required only 1–2 h of students' time and yielded increased grades on course quizzes and exams, although the intervention designed by Vemu et al. (2022), which only took 15-min to complete, did not yield a statistically significant increase in strategy use. Combined with the findings of the present study, it is unclear the extent to which an intervention as brief as 15 min can yield meaningful change, despite the potential value of such brief interventions.

Measures of strategy use

In the present study, we aimed to gather students' strategy use via complementary measures of strategy use (i.e., descriptions of their studying and CAMS subscales). Results for the two measures differed in that increases were evidenced on the former, more proximal, measure of the specific strategies described in the module, as well as some of the more distal measures (i.e., two of the four subscales). Specifically, while the descriptions of students' study strategies were coded based on the strategies discussed in the module; the CAMS subscales were less directly aligned. For example, none of the targeted strategies explicitly targeted critical thinking, while the elaboration strategy directly aligned with the elaboration subscale. Likewise, metacognitive self-regulation, which involved planning, monitoring, and regulating, loosely aligned with multiple strategies (e.g., retrieval practice and spaced practice) and the overarching aim of the video. Moreover, strategy use is not all or nothing (Sebesta and Bray Speth, 2017). Multiple factors contribute to the effectiveness of strategy use and the impact on learning (e.g., how long, often, or correctly a strategy was used). Walck-Shannon et al. (2021) accounted for this by looking at not only the number of strategies students used when studying for the exam, but also the proportion of time they studied using each of the active strategies. However, even in that approach, it was unclear how deeply or correctly each strategy was used (e.g., superficially or in ways that align with best practices). Consequently, our measure of students' self-reported strategy use was limited in the fact that we were not able to gauge the extent to which they used the respective strategy (e.g., just once or frequently) or how effectively they used it (e.g., in line with best practices or not).

In contrast with other interventions, our results revealed no differences for two of the CAMS subscales (i.e., elaboration and organization). Sebesta and Bray Speth (2017), for example, found that after their intervention, organizational strategies increased (i.e., keeping records, goal setting and planning, and reviewing graded work). Kritzinger et al. (2018) studied differences between students identified as at-risk, higher performing, as well as those in the "murky middle" (p. 2). They found differences between how students in each of these groups studied. For example, higher performing students were more likely to use metacognitive self-regulation and elaboration. This is particularly notable given that certain strategies can evoke desirable difficulties and may be challenging for students to utilize without additional support (Kritzinger et al., 2018; Walck-Shannon et al., 2021). Indeed, students may struggle or resist adopting new study strategies without being explicitly taught how to utilize them in their own context (Vemu et al., 2022).

Supporting all students' learning

An increase in strategy use ultimately only matters if concomitant changes are evidenced with regard to students' learning and performance outcomes. One of the key contributions of this study was that there was, in effect, a simple effect of changes in students' described strategy use on changes in exam score, in line with similar findings of both Walck-Shannon et al. (2021) and Bernacki et al. (2020).

Yet, students do not all enter college with the same level of preparation (Ley and Young, 1998), and students with lower GPAs may rely on more rehearsal-based study strategies than their peers with higher GPAs (Marbach-Ad et al., 2016). In contrast, students with higher GPAs may not recognize the need for using certain strategies or admit they used or need them (Kritzinger et al., 2018). A meta-analytic review by Credé and Phillips (2011) found some subscales to correlate with GPA (e.g., metacognitive self-regulation) but not others (e.g., elaboration or organization). Given this conflicting past research, we investigated the role of GPA, but we did not find any evidence that GPA moderated the strength of the relationship between changes in strategy use and exam score—students with higher GPAs did not have a stronger or weaker relationship than those with lower GPAs.

Limitations and areas for further exploration

Similar to Hensley et al. (2021), we also focused on growth from pretest to posttest, which allowed us to take into consideration students' extant strategy use and prior knowledge as well as the fact that exam 3 was more difficult than exam 2. While we intentionally provided all students interested in participating in the research with the brief, online module, to potentially support their learning, we recognize that without a control group, we cannot definitively attribute these changes to the module or make claims of causal inferences (see also Vemu et al., 2022). Of note, however, the sample of students who participated in the study did evidence less of a decrease in scores from exam 2 to exam 3 (i.e., 0.5%) compared to the overall average decrease in each section (i.e., between 2 and 3%), although such difference might be a result of selection bias. While it was not within the scope of the present study, we hope to see future research continue investigating the benefits of brief innovations to support students' study strategy use using experimental designs. Of note, there was nothing related to study techniques covered in the standard course instruction, and we have no knowledge of other interventions available to students in the class that would serve as an alternative explanation for the increase in strategies used, specific to those covered in the brief, online module.

Future research should also continue to explore the role of GPA. For feasibility reasons, it was not possible to obtain official student GPAs for this study, as such we gathered GPA via self-report. Discrepancies exist between self-reported GPA and official GPA. Indeed, Kuncel et al. (2005) found in their meta-analysis that, overall, self-reported GPA had increased error and decreased reliability, as individuals tend to have positive bias in their reporting, resulting in restricted range. However, after examining the moderation effects of various individual difference variables, the authors of the meta-analysis also forwarded an "ideal situation" where "more faith can be placed" on results derived from self-reported GPA (i.e., "self-reported grades from college students who have done well in school and have high cognitive ability scores," p. 76). We argue that while collecting GPA via self-report is less than ideal, our sample (i.e., high-achieving college students) is among those that have a more viable justification for use. While self-reported GPA may be artificially inflated, the moderately high overall nature of the sample suggests that the college students were overall high achieving. Taken together, there is a need to further investigate the impact of the module directly, using a comparison or control condition with a sample of students who have a wider range of GPA.

We hope to see continued research in this area to develop and evaluate more intensive interventions with reasonable time commitments and costs for both students and instructors. Using the present study as a case in point, students described using more effective strategies, although most still only used one or two of the strategies they learned about. This may have been related to the design of the module (i.e., students only personalized a plan for using their top two strategies) or the fact that most students selected to personalize a plan for strategies that were already commonly employed (e.g., retrieval practice). Thus, one future direction could be to extend the intervention by providing repeated (i.e., spaced) exposures to the video and allowing students to focus on different strategies each time. This would give students an opportunity to expand their repertoire of strategies and allow them to reflect on prior attempts and implement novel strategies. Alternatively, future research could explore other novel approaches to strategy interventions, for example, strategies that promote the use of collaborative or interactive strategies (e.g., forming study groups to promote small-group learning; Springer et al., 1999).

Finally, given the complexity of measuring strategy use, combined with the modest and somewhat mixed results of the present study, future research should continue to investigate ways to assess and gauge students' strategy use with these measures and others. Of note, one limitation of the present study is that we did not account for potential family-wise error (e.g., Bonferroni correction) in our analyses of the five CAMS subscales, which would be overly conservative in this context (e.g., sample size, number of subscales, and power). Interpretation of effect sizes, which align with a small effect for both critical thinking and metacognitive self-regulation, however, can serve as additional evidence that these results are less likely a result of Type 1 error. Additionally, the measures we employed did not assess the quality with which students used the various strategies (e.g., did they use them effectively) or the extent to which they used them (e.g., just once or in every study session). Future research should explore complementary measures that can better gauge the complexity of students' study strategy use.

Introductory human anatomy and physiology courses can serve as a "gateway" into health sciences careers, and it is a critical opportunity to examine interventions that can support students in this specific area. Supporting undergraduate biology students earlier in their academic pathways could positively impact the trajectory of their success in the field (Kritzinger et al., 2018). Even with a greater level of subjective value and motivation to succeed (Sullins et al., 1995), they may lack knowledge about SRL strategies and how to implement them (Sebesta and Bray Speth, 2017), limiting their learning and ultimately their success. Continued exploration is needed to recognize the challenges and constraints that students are navigating in order to identify and delineate brief, low-cost interventions that promote enhanced strategy use, learning, and academic performance for students in introductory human anatomy and physiology courses.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Arizona State University's Institutional Review Board. The patients/participants provided their written informed consent to participate in this study.

Author contributions

CF led the conception, design, and analysis of the study as well as manuscript writing. ES, AM, and LY contributed to the coding, analysis, and writing-up portions of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research project was supported by internal research grants awarded by both Mary Lou Fulton Teachers College and the Institute for Social Sciences Research at Arizona State University (ASU) as well as the ASU Open Access Publication Fund.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Adesope, O. O., Trevisan, D. A., and Sundararajan, N. (2017). Rethinking the use of tests: a meta-analysis of practice testing. *Rev. Educ. Res.* 87, 659–701. doi: 10.3102/0034654316689306

Ames, C., and Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *J. Educ. Psychol.* 80, 260–267. doi: 10.1037/0022-0663.80.3.260

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Benjamin, A. S., and Tullis, J. (2010). What makes distributed practice effective? *Cogn. Psychol.* 61, 228–247. doi: 10.1016/j.cogpsych.2010.05.004

Bernacki, M. L., Vosicka, L., and Utz, J. C. (2020). Can a brief, digital skill training intervention help undergraduates "learn to learn" and improve their STEM achievement? *J. Educ. Psychol.* 112, 765–781. doi: 10.1037/edu0000405

Bjork, E. L., and Bjork, R. A. (2011). "Making things hard on yourself, but in a good way: creating desirable difficulties to enhance learning" in *Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society*. eds. M. A. Gernsbacher, R. W. Pew, L. M. Hough and J. R. Pomerantz (New York, NY: Worth Publishers), 56–64.

Blasiman, R. N., Dunlosky, J., and Rawson, K. A. (2017). The what, how much, and when of study strategies: comparing intended versus actual study behaviour. *Memory* 25, 784–792. doi: 10.1080/09658211.2016.1221974

Broadbent, J., and Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: a systematic review. *Internet High. Educ.* 27, 1–13. doi: 10.1016/j.iheduc.2015.04.007

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 1118–1133. doi: 10.1037/a0019902

Cook, E., Kennedy, E., and McGuire, S. Y. (2013). Effect of teaching metacognitive learning strategies on performance in general chemistry courses. *J. Chem. Educ.* 90, 961–967. doi: 10.1021/ed300686h

Credé, M., and Phillips, L. A. (2011). A meta-analytic review of the motivated strategies for learning questionnaire. *Learn. Individ. Differ.* 21, 337–346. doi: 10.1016/j.lindif.2011.03.002

Dignath, C., and Büettner, G. (2008). Components of fostering self-regulated learning among students: a meta-analysis on intervention studies at primary and secondary school level. *Metacogn. Learn.* 3, 231–264. doi: 10.1007/s11409-008-9029-x

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., and Willingham, D. T. (2013). Improving students' learning with effective learning techniques: promising directions from cognitive and educational psychology. *Psychol. Sci. Public Interest* 14, 4–58. doi: 10.1177/1529100612453266

Eccles, J. S. (1983). "Expectancies, values, and academic behaviors" in *Achievement and Achievement Motives*. ed. J. T. Spence (San Francisco: Freeman), 75–146.

Hartwig, M. K., and Dunlosky, J. (2012). Study strategies of college students: are self-testing and scheduling related to achievement? *Psychon. Bull. Rev.* 19, 126–134. doi: 10.3758/s13423-011-0181-y

Hattie, J., Biggs, J., and Purdie, N. (1996). Effects of learning skills interventions on student learning: a meta-analysis. *Rev. Educ. Res.* 66, 99–136. doi: 10.3102/00346543066002099

Hayes, A. F. (2021). The PROCESS macro for SPSS and SAS (version 4.1). Available at: http://processmacro.org/index.html

Hensley, L., Kulesza, A., Peri, J., Brady, A. C., Wolters, C. A., Sovic, D., et al. (2021). Supporting undergraduate biology students' academic success: comparing two workshop interventions. *CBE Life Sci. Educ.* 20:ar60. doi: 10.1187/cbe.21-03-0068

Karpicke, J. D., Butler, A. C., and Roediger, H. L. (2009). Metacognitive strategies in student learning: do students practise retrieval when they study on their own? *Memory* 17, 471–479. doi: 10.1080/09658210802647009

Kiewra, K. A. (2002). How classroom teachers can help students learn and teach them how to learn. *Theory Pract.* 41, 71–80. doi: 10.1207/s15430421tip4102_3

Koch, A. K. (2017). "It's about the gateway courses: defining and contextualizing the issue" in *Improving Teaching, Learning, Equity, and Success in Gateway Courses. New Directions for Higher Education*. ed. A. K. Koch, vol. *180* (Hoboken, New Jersey: Jossey-Bass), 11–17.

Kornell, N., and Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychon. Bull. Rev.* 14, 219–224. doi: 10.3758/BF03194055

Kritzinger, A., Lemmens, J.-C., and Potgieter, M. (2018). Learning strategies for first-year biology: toward moving the murky middle. *CBE Life Sci. Educ.* 17, 1–13. doi: 10.1187/cbe.17-10-0211

Kuncel, N. R., Credé, M., and Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: a meta-analysis and review of the literature. *Rev. Educ. Res.* 75, 63–82. doi: 10.3102/00346543075001063

Ley, K., and Young, D. B. (1998). Self-regulation behaviors in underprepared (developmental) and regular admission college students. *Contemp. Educ. Psychol.* 23, 42–64. doi: 10.1006/ceps.1997.0956

Lopez, E. J., Nandagopal, K., Shavelson, R. J., Szu, E., and Penn, J. (2013). Self-regulated learning study strategies and academic performance in undergraduate organic chemistry: an investigation examining ethnically diverse students. *J. Res. Sci. Teach.* 50, 660–676. doi: 10.1002/tea.21095

Marbach-Ad, G., Rietschel, C., and Thompson, K. V. (2016). Validation and application of the survey of teaching beliefs and practices for undergraduates (STEP-U): identifying factors associated with valuing important workplace skills among biology students. *CBE Life Sci. Educ.* 15:ar59. doi: 10.1187/cbe.16-05-0164

Mayer, R. E., and Anderson, R. B. (1992). The instructive animation: helping students build connections between words and pictures in multimedia learning. *J. Educ. Psychol.* 84, 444–452. doi: 10.1037/0022-0663.84.4.444

McDaniel, M. A., and Donnelly, C. M. (1996). Learning with analogy and elaborative interrogation. *J. Educ. Psychol.* 88, 508–519. doi: 10.1037/0022-0663.88.3.508

Memorize Academy [Username] in collaboration with the Learning Scientists (2016). How to study effectively for school or college [top 6 science-based study skills] [video]. Youtube. Available at: https://www.youtube.com/watch?v=CPxSzxylRCI

Minchella, D. J., Yazvac, C. W., Fodrea, R. A., and Ball, G. (2002). Biology resource seminar: first aid for the first year. *Am. Biol. Teach.* 64, 352–357. doi: 10.2307/4451310

Muteti, C. Z., Zarraga, C., Jacob, B. I., Mwarumba, T. M., Nkhata, D. B., Mwavita, M., et al. (2021). I realized what I was doing was not working: the influence of explicit teaching of metacognition on students' study strategies in a general chemistry I course. *Chem. Educ. Res. Pract.* 22, 122–135. doi: 10.1039/d0rp00217h

Pintrich, P. R., Smith, D. A. F., Garcia, T., and McKeachie, W. J. (1991). A manual for the use of the motivated strategies for learning questionnaire (MSLQ), The University of Michigan, Ann Arbor.

Pomerance, L., Greenberg, J., and Walsh, K. (2016). Learning about learning: what every new teacher needs to know. National Council on Teacher Quality. Available at: https://www.nctq.org/publications/Learning-About-Learning:-What-Every-New-Teacher-Needs-to-Know

Rawson, K. A., Thomas, R. C., and Jacoby, L. L. (2014). The power of examples: illustrative examples enhance conceptual learning of declarative concepts. *Educ. Psychol. Rev.* 27, 483–504. doi: 10.1007/s10648-014-9273-3

Roediger, H. L., and Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends Cogn. Sci.* 15, 20–27. doi: 10.1016/j.tics.2010.09.003

Roediger, H. L., Putnam, A. L., and Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. *Psychol. Learn. Motiv.* 55, 1–36. doi: 10.1016/B978-0-12-387691-1.00001-6

Rohrer, D. (2012). Interleaving helps students distinguish among similar concepts. *Educ. Psychol. Rev.* 24, 355–367. doi: 10.1007/s10648-012-9201-3

Schneider, M., and Preckel, F. (2017). Variables associated with achievement in higher education: a systematic review of meta-analyses. *Psychol. Bull.* 143, 565–600. doi: 10.1037/bul0000098

Sebesta, A. J., and Bray Speth, E. (2017). How should I study for the exam? Self-regulated learning strategies and achievement in introductory biology. *CBE Life Sci. Educ.* 16:ar30. doi: 10.1187/cbe.16-09-0269

Springer, L., Stanne, M. E., and Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: a meta-analysis. *Rev. Educ. Res.* 69, 21–51. doi: 10.3102/00346543069001021

Sullins, E. S., Hernandez, D., Fuller, C., and Tashiro, J. S. (1995). Predicting who will major in a science discipline: expectancy–value theory as part of an ecological model for studying academic communities. *J. Res. Sci. Teach.* 32, 99–119. doi: 10.1002/tea.3660320109

Theobald, M. (2021). Self-regulated learning training programs enhance university students' academic performance, self-regulated learning strategies, and motivation: a meta-analysis. *Contemp. Educ. Psychol.* 66:101976. doi: 10.1016/j.cedpsych.2021.101976

Vemu, S., Denaro, K., Sato, B. K., Fisher, M. R., and Williams, A. E. (2022). Moving the needle: evidence of an effective study strategy intervention in a community college biology course. *CBE Life Sci. Educ.* 21:ar24. doi: 10.1187/cbe.21-08-0216

Walck-Shannon, E. M., Rowell, S. F., and Frey, R. F. (2021). To what extent do study habits relate to performance? *CBE Life Sci. Educ.* 20:ar6. doi: 10.1187/cbe.20-05-0091

Weinstein, Y., McDermott, K. B., and Roediger, H. L. (2010). A comparison of study strategies for passages: rereading, answering questions, and generating questions. *J. Exp. Psychol. Appl.* 16, 308–316. doi: 10.1037/a0020992

Weinstein, Y., Smith, M., and Caviglioli, O. (2023a) Our answers to frequently asked questions. Available at: https://www.learningscientists.org/faq

Weinstein, Y., Smith, M., and Caviglioli, O. (2023b) Six strategies for effective learning: materials for teachers and students. Available at: https://www.learningscientists.org/downloadable-materials

Wingate, U. (2007). A framework for transition: supporting 'learning to learn' in higher education. *High. Educ. Q.* 61, 391–405. doi: 10.1111/j.1468-2273.2007.00361.x

Zhao, N., Wardeska, J. G., McGuire, S. Y., and Cook, E. (2014). Metacognition: an effective tool to promote success in college science learning. *J. Coll. Sci. Teach.* 43, 48–54. doi: 10.2505/4/jcst14_043_04_48

Zimmerman, B. J. (2000). "Attaining self-regulation: a social cognitive perspective" in *Handbook of Self-Regulation*. eds. M. Boekarts, P. R. Pintrich and M. Zeidner (San Diego, CA: Academic Press), 13–39.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: an overview. *Theory Pract.* 41, 64–70. doi: 10.1207/s15430421tip4102_2

Zimmerman, B. J., and Schunk, D. H. (Eds.) (1989). *Self-Regulated Learning and Academic Achievement: Theory, Research, and Practice*. New York, NY: Springer.

frontiers | Frontiers in Education

TYPE Curriculum, Instruction, and Pedagogy
PUBLISHED 02 June 2023
DOI 10.3389/feduc.2023.1130865

OPEN ACCESS

EDITED BY
Wang-Kin Chiu,
The Hong Kong Polytechnic University, China

REVIEWED BY
Evelyn Valdez-Ward,
University of Rhode Island, United States
Yu-Chun Kuo,
Rowan University, United States

*CORRESPONDENCE
Manyu Li
✉ manyu.li@louisiana.edu

RECEIVED 23 December 2022
ACCEPTED 09 May 2023
PUBLISHED 02 June 2023

CITATION
Li M (2023) Teaching beginner-level computational social science: interactive open education resources with learnr and shiny apps. *Front. Educ.* 8:1130865.
doi: 10.3389/feduc.2023.1130865

Teaching beginner-level computational social science: interactive open education resources with learnr and shiny apps

Manyu Li*

Department of Psychology, University of Louisiana at Lafayette, Lafayette, LA, United States

The article presents the use of open, inclusive pedagogical frameworks to develop a series of open education resources (OERs), specifically, interactive shiny web applications for teaching beginner-level computational social science (CSS) in undergraduate social science education. CSS is an emerging discipline that integrates data science methods with social science theories and research designs. However, undergraduate social science students could face a lack of a sense of STEM identity or belonging. Also, compared to STEM majors in natural science or engineering, average undergraduate social science students come in with limited mathematical or statistical knowledge. The OERs developed and tested in this article are designed with pedagogical frameworks that help overcome these barriers faced by students from diverse backgrounds and offer students a jump-start in learning CSS. This article presents the details of the tools, classroom implementation (in the form of a 6-week workshop series), the pedagogy frameworks applied, and the assessment methods and outcomes.

KEYWORDS

computational social science, R programming, social science, research methods, pedagogy, shiny apps, undergraduate education, science technology engineering mathematics (STEM)

1. Introduction

Computational Social Science (CSS), formalized and popularized in the phenomenal paper by Lazer et al. (2009), is an emerging and evolving discipline defined as "the development and application of computational methods to complex, typically large-scale, human (sometimes simulated) behavioral data" (Lazer et al., 2020, p. 1060). In a little over a decade of development, CSS has attracted support and investments from top-tier institutions and the industry (Kim, 2021), creating a workforce with high demand in the tech industry, public policy decision-making, etc. The rise of CSS calls for a "paradigm for training new scholars" (Lazer et al., 2009, p. 722). Developing instructional methods for CSS education in social science is thus needed to improve the field.

This article aims at contributing to the CSS education literature by introducing instructional innovations, paired with open and inclusive pedagogical frameworks, that can be used to connect social science novices to the world of CSS learning. Specifically, to help students overcome challenges to step into the CSS world, a series of openly-available Rapport-building, Equitable, Learner-centered, Authentic Computational Social Science (RELACSS) web applications written in R language was developed and implemented in undergraduate social science education settings.

1.1. Preparing students to enter the CSS curriculum

Training "computationally literate social scientists" (Lazer et al., 2009, p. 722) may face several challenges. Firstly, undergraduate social science students were found to often lack a sense of STEM identity or belonging (Berndt et al., 2021; Esnard et al., 2021). This lack of STEM identity may contribute to barriers to learning computation-related skills (Berndt et al., 2021; Esnard et al., 2021). For example, their low sense of belonging/identity in computing (Chew and Dillon, 2014; Davies et al., 2015; Lawton and Taylor, 2020) may result in attentional biases (Okon-Singer, 2018; Cui et al., 2019) that further lower their engagement and motivation (Lawton and Taylor, 2020) and prohibit them from moving on to the next level (Davies et al., 2015). Additionally, these challenges have been identified among the under-resourced, marginalized under-represented minoritized (URM) population in STEM (Lisberg and Woods, 2018; Singer et al., 2020). Non-inclusive practices in the teaching process, stereotypes communicated, and perceived stereotype threats were suggested as potential reasons behind the achievement and belonging gap in URMs (Steele et al., 2002; Walton et al., 2015; Rattan et al., 2018; Liu et al., 2021; Yeager et al., 2022).

Second, compared to STEM majors in natural science or engineering, average undergraduate social science students come in with limited mathematical or statistical knowledge (Berndt et al., 2021) and thus find data science challenging (Dong et al., 2020). The lack of a foundation may have contributed to students' low motivation to learn and to advance in subjects that require prior knowledge in computation (Davies et al., 2015; Lawton and Taylor, 2020). Further, many do not expect to use statistics/computation to analyze data when choosing social sciences as their major (Esnard et al., 2021), nor do they expect to use the knowledge after graduation (Berndt et al., 2021). Thus, despite the availability of some free beginner materials online, students may not have the motivation to approach the materials. Even if students attempt to access the free resources available in data science or CSS (e.g., YouTube videos, free Massive Online Open Courses, etc.), complete novices may find the materials to be challenging and may even feel confused when accessing the materials. Therefore, the materials, even though they are freely available, may hinder students from exploring their interest in CSS.

RELACSS is designed to address these challenges and to increase students' interest in pursuing the CSS curriculum. First, in terms of content knowledge, RELACSS web apps focused on the very basics of what data analysis and coding are in social science. The goal is not to include all CSS topics, but to create interesting and approachable introductory CSS materials for complete novices. It is designed to reshape students' beliefs that data science is not for them and to increase their interest in further pursuing CSS. Second, as further detailed in the pedagogical frameworks, RELACSS web apps were designed with scaffolding and learner-centered principles to help students relate to the materials. Also, the web apps were built on social science students' background knowledge, rather than computer science; therefore, students are less likely to feel distant or out of place when learning the materials. Finally, RELACSS web apps are all built as interactive open education resources (OER)[1]. in the form of Shiny web applications. The interactive nature of the web apps helps students visualize and interact with the codes and materials without being overwhelmed with math or statistics symbols. The OER nature of the materials makes accessing the materials affordable (free) for all students and instructors who may adapt the materials to their courses.

1.2. Using multicultural data to create adaptable OERs and open assignments

Part of the lesson design for instructors from CSS or related fields is to identify data topics or data sets to which students can relate. Identifying topics or data examples that students relate to increases the personal relevance and authenticity of the learning materials, which in turn, increases students' learning outcomes (Lee et al., 2021). RELACSS addresses this by using multicultural data examples, allowing instructors from different backgrounds to adapt the materials and further contextualize the materials for their unique student population.

The RELACSS lessons developed in this study include data examples created using the World Value Survey data. The World Value Survey is an open data project with data collected from up to 80 countries around the world (Inglehart et al., 2022). The survey contains various questions on values and attitudes in work, family, politics, etc. The data were used throughout the lessons, including data examples, practice exercises, and a final project. Therefore, instructors can choose variables and countries that fit their learning contexts.

Using open data with a large number of variables and sub-samples (e.g., individuals nested within countries) also helps to create open assignments. Open assignments or OER-enabled assignments are assignments that meet the 5R principle of OERs (Wiley and Hilton, 2018). The 5R principle states that for materials to be OERs, they must allow the public to freely reuse, retain, revise, remix, and redistribute. Open assignments build upon this to encourage students to create renewable assignments that use OERs and then share with the 5R principle of OERs (Wiley and Hilton, 2018). In other words, instead of creating non-sustainable assignments that students discard after finishing the course, open assignments encourage students to take ownership by creating their own works and sharing for future learners openly (Wiley and Hilton, 2018; Allen and Katz, 2019). This type of open pedagogy also helps generate diverse, localized examples for future use (Ryan, 2022). In the final project built into the RELACSS learning experiences, students are asked to generate a new research question and conduct data analyses. Instead of discarding the assignments after the workshop, students share them openly.

1 Open Education Materials (OER) are defined as educational materials that follow the 5R principle, i.e., allowing users to freely reuse, retain, revise, remix, and redistribute (Wiley et al., 2013).

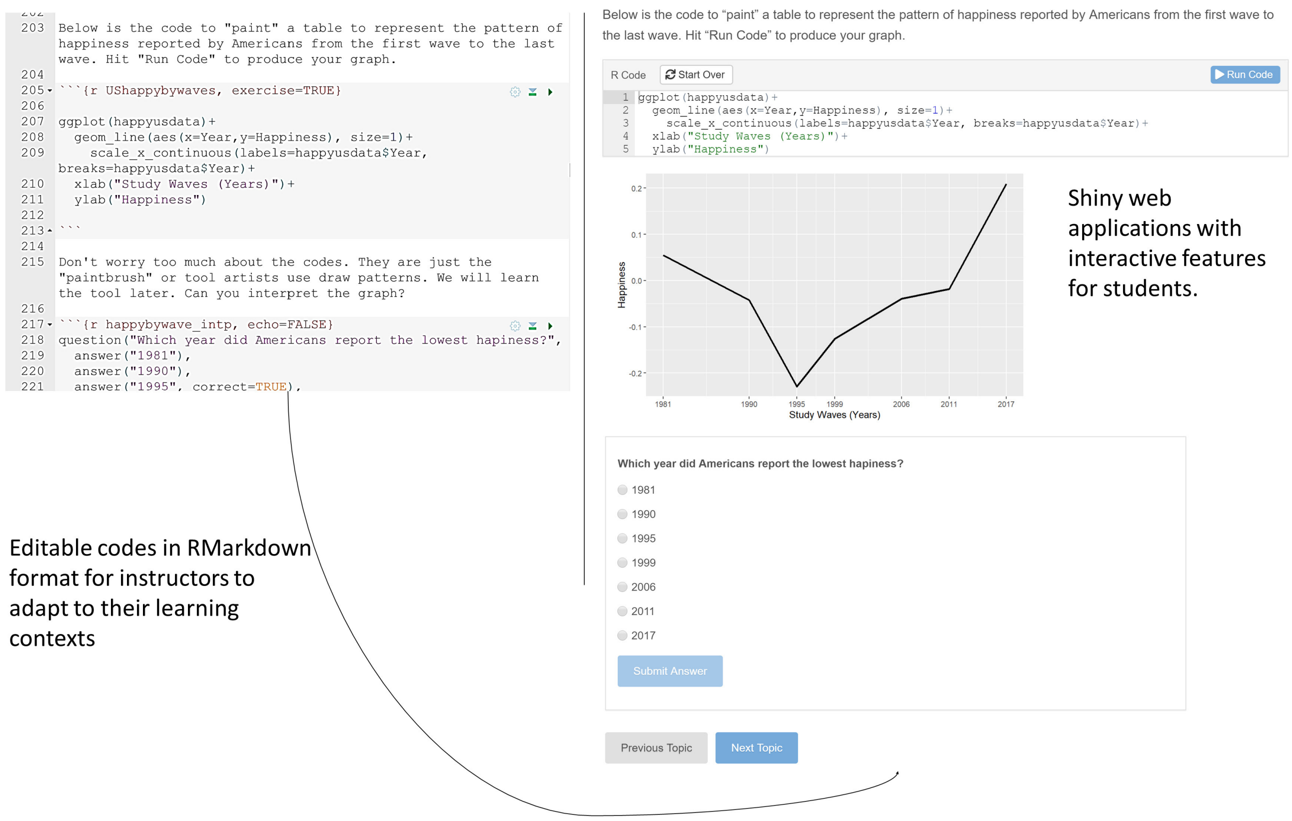

FIGURE 1
Source code **(left)** of the interactive shiny web applications **(right)** can be edited by instructors to teach in unique learning contexts.

1.3. Using Shiny and Learnr to create adaptable interactive OERs

Interactive features are crucial in overcoming students' challenges. Many previous studies in STEM learning have demonstrated that interactivity is crucial in sustaining students' interest and attention in the materials and to improve performance (Lindgren and DeLiema, 2022). For teaching topics like coding, it is important to have a comprehensive environment where teaching materials, examples, data, codes, and practices are all in one place. The teaching and learning experiences can also flow smoothly when students do not have to shift from various sources (platforms, books, software). To achieve these purposes, RELACSS lessons are built on two R packages, *shiny* and *learnr*, and are launched on a single website.

Shiny is an R package that allows app developers to design interactive web applications using R languages (Chang et al., 2022). The web applications can be deployed locally or on a server (e.g., the shiny apps server) and shared with others. Shiny has been used for creating educational materials in various STEM subjects, such as ecology (Moore et al., 2022), statistics (Doi et al., 2016), and biology (Weigelt et al., 2021). The package, *learnr* was developed to increase efficiency in writing interactive shiny apps for education purposes (Aden-Buie et al., 2022).

Shiny web applications, being open-sourced, are great candidates for creating interactive OERs. Developers can easily share source codes for app users to not only retain and reuse, but also to revise, remix, and redistribute (i.e., the 5Rs of OERs). For example, as shown in Figure 1, in the introductory materials developed, the World Value Survey data (Inglehart et al., 2022) were used to demonstrate how to understand happiness ratings among US people in the past 40 years. Instructors from another country can edit the source codes and change the target country or target variables. Instructors may also add practice questions and charts by editing the codes.

In summary, to overcome the challenges of social science novice students in starting to learn computational knowledge, RELACSS beginner OERs were developed with interactive and open features. Several pedagogical frameworks were then applied in designing the learning objectives, contents, and in-class activities. RELACSS was then implemented as a 6-week workshop for university students majoring in social sciences.

2. Pedagogical goals and frameworks

As explained in Section 1, the ultimate pedagogical goal of RELACSS is to prepare students to enter the CSS curriculum by overcoming students' challenges, mainly, (1) creating learner-centered, authentic (approachable) content, (2) cultivating a rapport-building, inclusive climate to increase students' STEM identity or sense of STEM belonging, and (3) generating open and interactive materials to make CSS affordable and interesting to students. Several pedagogical frameworks were employed to achieve the goals.

2.1. Scaffolding: learner-centered, authentic learning experiences

Scaffolding is an instructional strategy (often computer-based) to guide students in building new knowledge systematically (Kim et al., 2017). Scaffolding has been suggested to be effective in helping students master skills or learn materials that they would otherwise not be able to do by themselves (Sherin et al., 2004; Holmes et al., 2013). Scaffolding requires instructors to be learner-centered and to focus on the current state of the students. Specifically, scaffolding requires instructors to identify students' zone of proximal development (ZPD), i.e., the distance between what learners cannot do on their own and what they can do on their own (Vygotskii, 1978; Wertsch, 1984). Scaffolding is realized in three ways in the design of RELACSS.

First, the content knowledge is taught progressively, starting with content that students can easily interpret and scaffold new knowledge. As previously mentioned, average undergraduate social science students often enter the university with limited mathematical and statistical knowledge to sufficiently understand beginner computation (Dong et al., 2020; Berndt et al., 2021). Students may find it difficult to step right into analyzing data and coding. Therefore, the content knowledge taught in RELACSS begins with describing and visualizing data using basic knowledge that is commonly taught in high schools, such as examining trends of a variable in a line graph or histogram. Further, RELACSS is designed for students who are complete novices. Donoho (2017) also suggested that beginner data science education should focus on teaching data gathering, preparation, and exploration (GPE) over modeling at the beginner level. Therefore, RELACSS materials focus on using R codes to learn basic GPE content, including reading data into R, data structure and variables types, exploratory research questions, data cleaning, data visualization, descriptive data analysis, and interpreting and communicating research results.

Second, the overall RELACSS beginner materials follow a progressive four-module structure of Bird-eye view, Use, Modify, and Create (BUMC). In the first module, *Bird-eye view*, students are exposed to a data story (Lee et al., 2015) that contains a sample "end-product", i.e., an engaging narrative and visualization explaining a data example. The design of this task is uniquely catered to social science students' lack of interest in data science (Berndt et al., 2021; Esnard et al., 2021) and to motivate students to pursue the next steps. The second to fourth modules follow the Use-Modify-Create structure (Lee et al., 2011) to increase the authenticity of the beginner computation exploration (Franklin et al., 2020; Weintrop, 2021). Specifically, in the second module, *Use*, students use existing materials and codes to learn basic data analytic skills. Then, the third module, *Modify*, invites students to consider how they can modify the RQs based on their interests and real-life experiences. In this stage, students modify the analytic plan and R codes and share their modifications with the class. Finally, the last module, *Create*, challenges students to collaboratively come up with new RQs and write a basic analytic plan and codes from scratch. As detailed in the next section, in addition to the cross-cultural tasks, each module will contain group reflections and co-design discussions to enhance equitable social ecologies in the learning process of RELACSS.

Third, RELACSS materials use social science research questions and data examples to build on social science students' existing knowledge and add new knowledge step-by-step. For example, students are guided to describe and visualize the trend of reported happiness among people in the US. Happiness or life satisfaction, being a topic examined across social science disciplines, can easily connect to social science students. After interpreting basic visualization, students then discuss and come up with factors relating to happiness to test. The analytic concepts and coding methods are embedded in these examples. Through these approachable social science analytic examples, students can associate new knowledge with their existing knowledge.

2.2. Rapport-building, equitable collaborative learning

Peer support is a key feature of RELACSS. As previously mentioned, because of stereotype threats and other associated factors, students, particularly URMs, may feel out of place when learning STEM subjects, resulting in a low sense of belonging and STEM identity (Sax et al., 2018). Collaborative learning can help cultivate an inclusive classroom climate (Nishina et al., 2019) which is crucial to students' success and sense of belonging (Rodriguez and Blaney, 2021; Goering et al., 2022). Collaborative learning has also been consistently found to improve student learning processes and outcomes (Echeverria et al., 2019; Micari and Pazos, 2020).

In the design of RELACSS, in-class discussion questions are embedded in the materials. Further, during the "Modify" stage, two to three students are grouped to modify research questions and codes to conduct analyses using the guided prompt embedded in the lessons. Finally, at the end of each class meeting, 5–10 min are reserved for group reflection. Students first fill out a brief form asking them to summarize their knowledge gained and the challenges experienced. Students shared their responses with one another and collaboratively reflected on the learning experiences. During the workshop, students were also asked to make design suggestions that enable them to learn the materials more effectively. The group reflections and co-design discussions both serve to enhance two-way communications and equitable social ecologies in the learning process of RELACSS (Gutiérrez and Jurow, 2016).

2.3. OER-enabled pedagogy

As briefly mentioned in Section 1.2, OER-enabled pedagogy (Wiley and Hilton, 2018) is employed in the "Create" module. Open or OER-enabled pedagogy can be referred to as the design of

TABLE 1 Module-level learning objectives.

Module		Module-level learning objectives
1 Bird-eye view	1.1	Describe what data analysis is
	1.2	State how social scientist solves RQs through data analysis
	1.3	Identify the role of data analysis in social science
2 Use	2.1	Describe fundamental steps of research process and data analysis
	2.2	Identify data structure through the use of cross-cultural open data
	2.3	Interpret the meaning of existing RQs
	2.4	Run existing R codes to solve RQs
	2.5	Interpret the results of the analysis performed
3 Modify	3.1	Modify RQ based on learners' interest in the cross-cultural open data
	3.2	Compare basic R codes and distinguish codes based on analytic purposes;
	3.3	Repeat the data analysis process from Module 2 with modified RQs
	3.4	Interpret the results of the analysis performed
4 Create	4.1	Design new RQs from the real-world cross-cultural open data
	4.2	Formulate a data analysis plan to answer the RQs developed
	4.3	Write and run R codes to administer the data analysis plan
	4.4	Present the findings and interpretation of the results

RQ = Research Question.

renewable or sustainable assignments or classroom activities using OERs (Wiley and Hilton, 2018). In the "Create" module, learners are asked to use the cross-cultural World Value Survey open data and the RELACSS OERs learned in the previous class meetings to generate a new research question, analytic plan, and codes. Then, students share the end product openly as a shiny web application, allowing other learners to access and learn from their analytic products.

2.4. Learning objectives

The design of RELACSS beginner OERs is guided by the following course-level learning objectives:

- Identify basic social science data analytic steps;
- Explore a real-world, open cross-cultural dataset and formulate testable research questions;
- Write simple R codes to execute data analytic tasks;
- Interpret and present data analysis results.

The module-level learning objectives can be found in Table 1. The contents of RELACSS materials can be found on the RELACSS instructional site (bit.ly/RELACSSweb). R language was chosen as the coding language taught in RELACSS because R is the most common coding language among social scientists (Eiler et al., 2020; Vance, 2021).

3. Method of implementation

3.1. Learning environment

RELACSS beginner materials were implemented as a 6-week, non-credit workshop for social science students in a university setting. None of the students had prior experience in analyzing data using R codes. Each week, the students met for one hour as a group. A total of 16 students spread across two groups completed the workshop. Students ranged from first-year undergraduate students to graduate students. Students were invited to bring their laptops, but computers were also provided for the students to learn. There were no technical requirements for the computers except that the device must be able to use any web browser to access the lessons (Chrome, Firefox, etc.). The author who is a faculty member in psychology taught the lessons.

During each session, students arrived at the classroom and prepared the lesson by opening the URL of the shiny app lessons. Students were first guided to read the learning objectives of the lesson (the first section of every lesson). The instructor then guided the students to go over the interactive shiny web applications. Embedded prompts were provided to students for participation in peer discussions. Students also had the opportunity to run the codes, respond to the interactive practice questions, and discuss the concepts learned.

The first week was the "Bird-eye View" module, inviting students to examine a data example. During the second to the third week, students applied existing codes and examples to learn the knowledge (the "Use" module). In the fourth week, students collaboratively modified the codes (the "Modify" module). Finally, during the last 2 weeks, students engaged in the "Create" module and complete the final project during class time. Students were encouraged to ask questions and discuss with peers when they create their final project products.

3.2. Assessment

Assessments were conducted according to the pedagogical goals of RELACSS. Student learning outcomes were measured in the format of a final project. The final project was assessed on the accuracy of the data analytic steps and codes, as well as the level of creativity. Students' CSS skills were scored using an adapted version of the American Statistical Association (ASA) Project Competition Rubric (ASA, 2022). The grading categories include (1) research question quality (clearly stated research questions and alignment with analytic methods), (2) raw data management, (3) data visualization, (4) data analysis, (5) conclusion (e.g., clear answers to research questions), and (6) overall presentation.

Additionally, after the workshop, students were assessed on their perception of the learning process. The assessments were developed based on the challenges the tools aimed to address (see Section 1) and the pedagogy goals. The assessment questions (AQ) and the associated assessment tools included

- AQ1. What is students' perception of interactivity and authenticity (i.e., being approachable and relatable) of the RELACSS OERs?

- AQ2. What is students' perception of the social aspects of the classroom experiences (rapport-building collaborative learning)?
- AQ3a. What are students' perceived changes before and after the lessons on their confidence in data analysis, interest in data analysis, knowledge of data analysis, and anxiety when thinking about data analysis?
- AQ3b. What are students' CSS identity, interests, and plans for pursuing data analysis/CSS further?

3.2.1. Materials and data analysis

The assessment tools, the final project rubric, along with the anonymized data can be found on the project's Open Science Foundation (OSF) site (https://osf.io/m87kh/). Details of the assessments are also reported along with the results in Section 4. The assessment and human research associated with the assessments were approved by the author's Institutional Review Board (IRB) under the approval number: IRB-22-123-PSYC-Gird. The data analysis for the assessment is primarily descriptive. Students' final project outcomes were summarized using the adapted version of the American Statistical Association (ASA) Project Competition Rubric (ASA, 2022) mentioned earlier. Then, to answer AQs, descriptive statistics (means and standard deviations) were computed and reported.

3.2.2. Study participants

Participation in the assessment portion is completely voluntary. A total of 10 students filled out the feedback form. Among the nine students who reported demographics, six identified as females, and three as males. Students were mostly URMs, including African/Black ($n = 5$), Middle-eastern ($n = 1$), and Mixed-race ($n = 1$). Only two identified as White. Students were spread among lower-class undergraduate students ($n = 1$), upper-class undergraduate students (junior or above; $n = 6$), and graduate students ($n = 2$).

4. Results and discussion

4.1. Final project outcomes

With the help of the instructor and peers, all students were able to produce their unique final project products. All final products can be viewed openly in the RELACSS instructional site (bit.ly/RELACSSweb). All students produced satisfactory final project products according to the rubric. Specifically, all students were able to formulate their unique research questions based on variables and countries available in the World Value Survey data. Ten out of the 16 students chose to compare groups or countries, and six students chose to examine relationships between chosen variables. All students produced an excellent level of research questions. In terms of data management codes, all students were able to take materials from previous RELACSS modules to guide them to write the codes. However, most students needed to ask questions in the process to clarify or confirm their coding accuracy. Students also tended to run redundant codes (lack of efficiency) or place codes in incorrect orders.

The biggest challenge in the final project creation appeared to be choosing the right visualization (e.g., box plots vs. line plots) and writing the visualization codes. This may be due to insufficient lesson time to go over a variety of plots. The in-class time was only sufficient for going over limited visualization examples. These examples did not necessarily match the types of research questions students are asking. Also, students were mostly unfamiliar with the principles of visualizing data based on variable types (e.g., differences in visualizing categorical and continuous variables). One way to resolve this is to limit the type of research questions students ask; alternatively, more time can be devoted to going over visualization rather than the codes.

Despite the challenges, with the help of the instructor, all students were able to identify the codes used to visualize their results. Students expressed that the visualization helps them write and present their interpretation of the results. This further underlines the importance of teaching visualization at the beginner level.

4.2. Students' perception of the learning process

In terms of students' perception of interactivity and authenticity of the RELACSS OERs (AQ1), students were asked to rate on a Likert scale from 0 to 10 their perception of whether the RELACSS web applications are authentic (relatable and approachable), learner-centered, informative, and interactive, as well as whether RELACSS helps them gain new knowledge. As shown in Figure 2, students reported high scores across all aspects. Specifically, students perceived RELACSS to be highly interactive ($M = 9.11$, $SD = 0.99$), learner-centered ($M = 8.89$, $SD = 0.99$), and authentic/relatable ($M = 9.11$, $SD = 0.99$). Students also perceived RELACSS to be informative ($M = 9.22$, $SD = 1.31$) and help gain new knowledge ($M = 8.44$, $SD = 1.07$).

Students' perception of the social aspects (rapport-building collaborative learning; AQ2) was assessed using one open-ended question, "Is there anything you like or dislike about the classroom learning experience?" Students, in general, responded positively to the classroom climate. Specifically, many students expressed that they felt bonded to other students (e.g., "felt connected to other students", "felt cared for and respected", "felt included and respected", etc.). Others mentioned that they "liked the group setting" and that "the group was small so all my questions were able to be answered and everyone was able to have individual help". One student mentioned that they wished to have more in-class time and suggested increasing class time and having homework assignments in the future.

Next, students were asked to rate their perceived change before and after the lessons (AQ3a) on their confidence in data analysis, interest in data analysis, knowledge of data analysis, and anxiety when thinking about data analysis using a 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree). "Data analysis" was chosen as the prompt because data analysis was the umbrella term

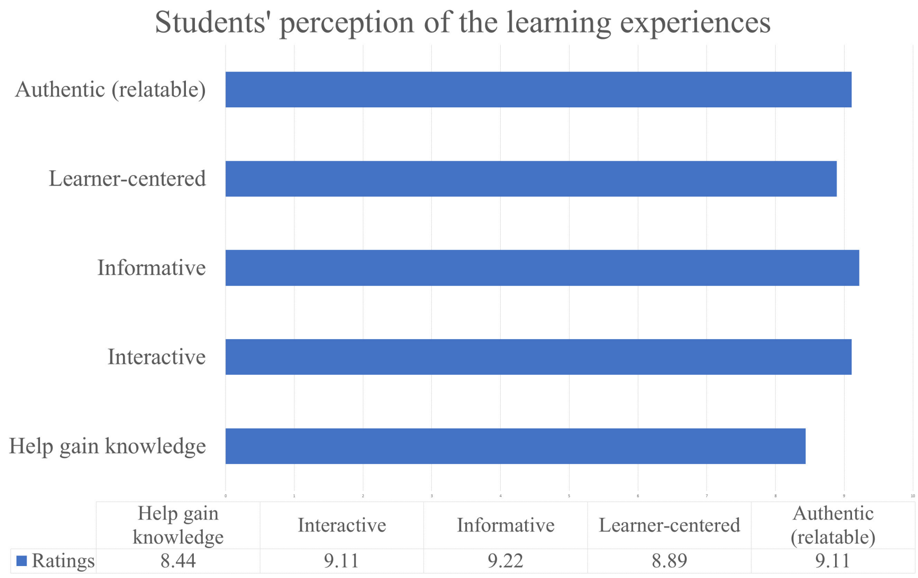

FIGURE 2
Average ratings of students' perception toward the RELACSS learning experiences.

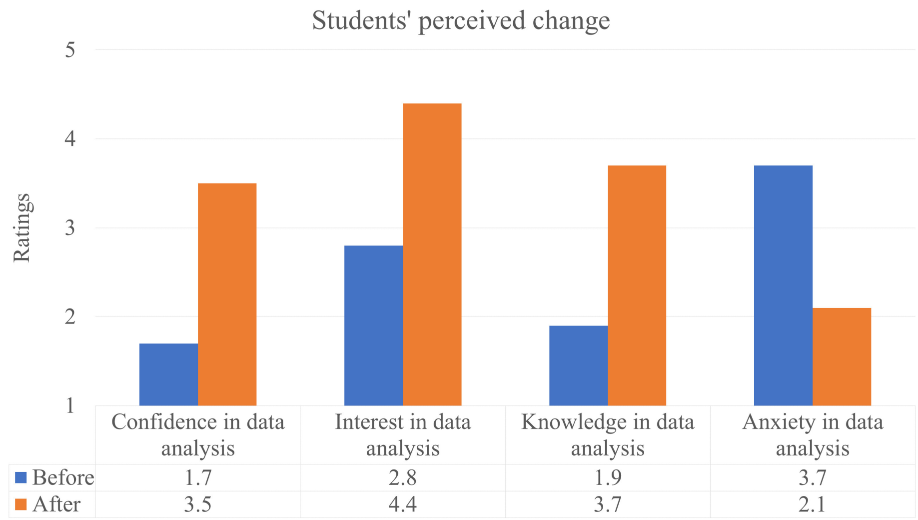

FIGURE 3
Average ratings of students' perception of change before and after the RELACSS learning experiences.

used to refer to the different aspects of knowledge taught during the implementation (e.g., coding, data management, data visualization, data description, etc.). As shown in Figure 3, students' average self-ratings on their confidence in data analysis ($M_{after-before} = 1.80$, $SD_{after-before} = 0.98$), interest in data analysis ($M_{after-before} = 1.60$, $SD_{after-before} = 1.11$), and knowledge in data analysis ($M_{after-before} = 1.80$, $SD_{after-before} = 0.87$) were about double after participating in RELACSS. Students' anxiety in data analysis also slightly dropped after the RELACSS experiences ($M_{after-before} = -1.60$, $SD_{after-before} = 0.80$). No inferential tests were performed due to a small assessment sample size.

Students' CSS identity, interests, and plans for pursuing data analysis/CSS in the future (AQ3b) were explored using multiple assessment questions. Students' CSS identity was assessed by asking students to rate from 0 to 10 their agreement on whether they see themselves as a future computational social scientist or data analyst. All students rated highly on the item ($M = 9.44$, $SD = 0.68$). Similarly, when using the same scale to report interest in further pursuing social science, students rated their interest highly ($M = 8.44$, $SD = 1.42$). Students were also asked if they would sign up for the next level of the CSS workshop in the following semester. All students were interested in signing up.

Finally, students' interest was further gauged using an open-ended question asking them in what ways the RELACSS experiences "has any impact on your future academic and career development or plan". Some students explicitly expressed that they became interested in a data science or CSS career and that the RELACSS experiences opened a new career possibility for them (e.g., "I would like to go into a career in data analysis", "The workshop definitely opened my eyes to career opportunities.", "I think I will be a lot more interested in doing data analysis and statistics for research studies during graduate school", "I am now more comfortable pursuing a career in data analysis related fields"). Students also shared that the RELACSS experiences changed how they saw their roles in data analysis (e.g., "This workshop made me feel like I have the ability to actually learn R and apply it in my future career, which I didn't think I had the capacity to do before.", "I would have been too scared to even consider data analysis as a career before the workshop").

4.3. Implications and lesson learned

The RELACSS open, interactive shiny web applications and the associated pedagogical frameworks are teaching innovations developed to address challenges novice social sciences may experience when learning CSS-related knowledge. The open nature of RELACSS will allow future instructors to adapt the web applications to their localized context. In addition to being open, the advantage of using the shiny applications for classroom instruction instead of having students work in R Studio or R console is that the tools do not require any software download or account registration/login. Instructors and students simply need to open the URLs for the shiny app and start typing codes and analyzing data using the embedded data examples. This saves time downloading software, opening the software, importing data, etc.

Results from the implementation showed that students felt positive toward the RELACSS experiences and agreed that RELACSS helped them gain new knowledge and move forward in pursuing the data science pathway further as social scientists. Specifically, students rated RELACSS to be highly interactive, learner-centered, and relatable. In terms of the social design of RELACSS, students expressed that they felt included during the experiences and felt comfortable joining the learning group weekly. Students found the instructional environment to be supportive, offering them individual help when needed. In addition, RELACSS appears to help students create a CSS/data analysis identity. Students shared that the RELACSS experiences helped them see what they could achieve in data analysis.

However, because of limited time, some aspects were not assessed. For example, students' perceived difficulties were not examined. Similarly, the effect of scaffolding was not fully examined. However, informal observations made during the 6-week interaction seemed to show that students were able to resolve challenges by asking questions during class, as they all managed to complete the tasks assigned during each lesson. Students' success in persisting to the end of the workshop and generating the final project end products (with in-class help) also supported that the level of difficulty appeared to be appropriate. Future researchers and educators may consider conducting a more thorough assessment, such as using formal classroom observations, interviewing students, or conducting focus groups. Researchers may also conduct design-based research to examine the different design/pedagogical components of the lessons. In addition, there was no quantitative assessment of students' perceived belonging and perception of classroom climate. The benefit of using merely a brief open-ended question for all social aspects is to allow students to freely share their perception of the classroom climate, particularly in aspects that the instructor may not notice. Future studies may consider using mixed methods to assess the social aspects.

One important lesson learned in the implementation process is that students may need more time to complete the RELACSS beginner materials. The implementation reported in this study was one hour per week for 6 weeks out of concerns about beginners' attention spans. Scheduling more than one hour outside everyone's class time was also a challenge. Some students expressed in the survey that they wanted more time to go deeper into the materials. These students tended to be lower-class undergraduate students. On the other hand, graduate students and some upper-class undergraduate students appeared to be comfortable with the current time setting. Future instructors (especially those who are teaching lower-class undergraduate students) may consider doing 1.5 h per week or extending the RELACSS experiences into an 8-week program. Another alternative is to incorporate the materials in regular 3-credit research methods or beginner computational social science courses. This option may fit better in institutions that are more open to creating a new course on CSS or updating current research methods or data analysis courses to include computational knowledge.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://osf.io/m87kh/.

Ethics statement

The studies involving human participants were reviewed and approved by Institutional Review Board, University of Louisiana at Lafayette. The patients/participants provided their written informed consent to participate in this study.

Author contributions

ML contributed to the conception, design, and assessment of the study and wrote, revised, read, and approved the manuscript.

Funding

This work was supported by the Open Education Resources (OER) Research Fellow awarded to the author by the Open Education Group and the Research Competitiveness Subprogram awarded to the author by the Board of Regents Louisiana [LEQSF(2019-22)-RD-A-23]. The Open Education Group supported the publication cost. The Board of Regents Louisiana supported the cost of the research activities and shiny web application development and deployment.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Aden-Buie, G., Schloerke, B., and Allaire, J. (2022). *learnr: Interactive Tutorials for R*. Available online at: https://rstudio.github.io/learnr/, https://github.com/rstudio/learnr

Allen, J. V., and Katz, S. (2019). Developing open practices in teacher education: an example of integrating OER and developing renewable assignments. *Open Prax*. 11, 311. doi: 10.5944/openpraxis.11.3.972

ASA (2022). *Scoring Rubric for American Statistical Association (ASA) Project Competition*. Available online at: https://www.amstat.org/asa/files/pdfs/EDU-ProjectJudgingRubric.pdf

Berndt, M., Schmidt, F. M., Sailer, M., Fischer, F., Fischer, M. R., and Zottmann, J. M. (2021). Investigating statistical literacy and scientific reasoning & argumentation in medical-, social sciences-, and economics students. *Learn. Individ. Differ*. 86, 101963. doi: 10.1016/j.lindif.2020.101963

Chang, W., Cheng, J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., et al. (2022). *shiny: Web Application Framework for R*. R package version 1.7.4.9000. Available online at: https://shiny.rstudio.com/

Chew, P. K. H., and Dillon, D. B. (2014). Statistics anxiety update: refining the construct and recommendations for a new research agenda. *Perspect. Psychol. Sci*. 9, 196–208. doi: 10.1177/1745691613518077

Cui, S., Zhang, J., Guan, D., Zhao, X., and Si, J. (2019). Antecedents of statistics anxiety: an integrated account. *Pers. Individ. Dif*. 144, 79–87. doi: 10.1016/j.paid.2019.02.036

Davies, E. L., Morys-Carter, W. L., and Paltoglou, A. E. (2015). Helping students to climb the mountain: a study to inform the development of a resource to improve the learning of statistics in psychology. *Psychol. Teach. Rev*. 21, 56–67. doi: 10.53841/bpsptr.2015.21.1.56

Doi, J., Potter, G., Wong, J., Alcaraz, I., and Chi, P. (2016). Web application teaching tools for statistics using r and shiny. *Technol. Innovat. Stat. Educ*. 9. doi: 10.5070/T591027492

Dong, A., Jong, M. S. Y., and King, R. B. (2020). How does prior knowledge influence learning engagement? The mediating roles of cognitive load and help-seeking. *Front. Psychol*. 11, 1–10. doi: 10.3389/fpsyg.2020.591203

Donoho, D. (2017). 50 years of data science. *J. Comput. Graph. Stat*. 26, 745–766. doi: 10.1080/10618600.2017.1384734

Echeverria, L., Cobos, R., and Morales, M. (2019). Improving the students computational thinking skills with collaborative learning techniques. *IEEE Rev. Iberoam. Tecnol. Aprend*. 14, 196–206. doi: 10.1109/RITA.2019.2952299

Eiler, B. A., Doyle, P. C., Al-Kire, R. L., and Wayment, H. A. (2020). Teaching computational social science skills to psychology students: an undergraduate research lab case study. *Scholar. Pract. Undergrad. Res*. 4, 5–14. doi: 10.18833/spur/4/1/5

Esnard, T. R., Alladin, F. M., and Samlal, K. C. (2021). Prior mathematics performance, statistics anxiety, self-efficacy and expectations for performance in statistics: a survey of social sciences students in a caribbean institution of higher education. *Stat. Educ. Res. J*. 20, 4. doi: 10.52041/serj.v20i1.98

Franklin, D., Coenraad, M., Palmer, J., Eatinger, D., Zipp, A., Anaya, M., et al. (2020). An analysis of use-modify-create pedagogical approach's success in balancing structure and student agency. In: *ICER 2020 - Proceedings of the 2020 ACM Conference on International Computing Education Research*. Virtual event, 14–24. doi: 10.1145/3372782.3406256

Goering, A. E., Resnick, C. E., Bradford, K. D., and Othus-Gault, S. M. (2022). Diversity by design: broadening participation through inclusive teaching. *New Direct. Commun. Coll*. 2022, 77–91. doi: 10.1002/cc.20525

Gutiérrez, K. D., and Jurow, A. S. (2016). Social design experiments: toward equity by design. *J. Learn. Sci*. 25, 565–598. doi: 10.1080/10508406.2016.1204548

Holmes, N. G., Day, J., Park, A. H. K., Bonn, D. A., and Roll, I. (2013). Making the failure more productive: scaffolding the invention process to improve inquiry behaviors and outcomes in invention activities. *Instruct. Sci*. 42, 523–538. doi: 10.1007/s11251-013-9300-7

Inglehart, R., Haerpfer, C., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., et al. (2022). *World Values Survey: All Rounds - Country-Pooled Datafile Version*. Technical report, JD Systems Institute. Available online at: https://www.worldvaluessurvey.org/WVSDocumentationWV7.jsp

Kim, B. (2021). Scaling up data science for the social sciences. *Harvard Data Sci. Rev*. 3. doi: 10.1162/99608f92.d3f14ea4

Kim, N. J., Belland, B. R., and Walker, A. E. (2017). Effectiveness of computer-based scaffolding in the context of problem-based learning for stem education: Bayesian meta-analysis. *Educ. Psychol. Rev*. 30, 397–429. doi: 10.1007/s10648-017-9419-1

Lawton, S., and Taylor, L. (2020). Student perceptions of engagement in an introductory statistics course. *J. Stat. Educ*. 28, 45–55. doi: 10.1080/10691898.2019.1704201

Lazer, D. M. J., Pentland, A., Adamic, L., Aral, S., Barabási, A.-, l., Brewer, D., et al. (2009). Computational social science. *Science* 323, 721–723. doi: 10.1126/science.1167742

Lazer, D. M. J., Pentland, A., Watts, D. J., Aral, S., Athey, S., Contractor, N., et al. (2020). Computational social science: obstacles and opportunities. *Science* 369, 1060–1062. doi: 10.1126/science.aaz8170

Lee, B., Riche, N. H., Isenberg, P., and Carpendale, S. (2015). More than telling a story: a closer look at the process of transforming data into visually shared stories. *IEEE Comp. Graph. Appl*. 35, 84–90. doi: 10.1109/MCG.2015.99

Lee, I., Martin, F., Denner, J., Coulter, B., Allan, W., Erickson, J., et al. (2011). Computational thinking for youth in practice. *ACM Inroads* 2, 32–37. doi: 10.1145/1929887.1929902

Lee, V. R., Wilkerson, M. H., and Lanouette, K. (2021). A call for a humanistic stance toward k–12 data science education. *Educ. Res*. 50, 664–672. doi: 10.3102/0013189X211048810

Lindgren, R., and DeLiema, D. (2022). Viewpoint, embodiment, and roles in STEM learning technologies. *Educ. Technol. Res. Dev*. 70, 1009–1034. doi: 10.1007/s11423-022-10101-3

Lisberg, A., and Woods, B. (2018). Mentorship, mindset and learning strategies: An integrative approach to increasing underrepresented minority student retention in a stem undergraduate program. *J. STEM Educ*. 19, 14–20.

Liu, S., Liu, P., Wang, M., and Zhang, B. (2021). Effectiveness of stereotype threat interventions: a meta-analytic review. *J. Appl. Psychol*. 106, 921–949. doi: 10.1037/apl0000770

Micari, M., and Pazos, P. (2020). Beyond grades: improving college students' social-cognitive outcomes in STEM through a collaborative learning environment. *Learn. Environ. Res*. 24, 123–136. doi: 10.1007/s10984-020-09325-y

Moore, T. N., Thomas, R. Q., Woelmer, W. M., and Carey, C. C. (2022). Integrating ecological forecasting into undergraduate ecology curricula with an r shiny application-based teaching module. *Forecasting* 4, 604–633. doi: 10.3390/forecast4030033

Nishina, A., Lewis, J. A., Bellmore, A., and Witkow, M. R. (2019). Ethnic diversity and inclusive school environments. *Educ. Psychol*. 54, 306–321. doi: 10.1080/00461520.2019.1633923

Okon-Singer, H. (2018). The role of attention bias to threat in anxiety: mechanisms, modulators and open questions. *Curr. Opin. Behav. Sci*. 19:26–30. doi: 10.1016/j.cobeha.2017.09.008

Rattan, A., Savani, K., Komarraju, M., Morrison, M. M., Boggs, C., and Ambady, N. (2018). Meta-lay theories of scientific potential drive underrepresented students' sense of belonging to science, technology, engineering, and mathematics (stem). *J. Pers. Soc. Psychol*. 115, 54–75. doi: 10.1037/pspi0000130

Rodriguez, S. L., and Blaney, J. M. (2021). “we’re the unicorns in STEM”: Understanding how academic and social experiences influence sense of belonging for latina undergraduate students. *J. Divers. High. Educ.* 14, 441–455. doi: 10.1037/dhe0000176

Ryan, M. (eds.) (2022). *Beyond the Traditional Essay: Increasing Student Agency in a Diverse Classroom With Nondisposable Assignments. Series in Education.* Wilmington, NC: Vernon Press.

Sax, L. J., Blaney, J. M., Lehman, K. J., Rodriguez, S. L., George, K. L., and Zavala, C. (2018). Sense of belonging in computing: the role of introductory courses for women and underrepresented minority students. *Soc. Sci.* 7, 122. doi: 10.3390/socsci7080122

Sherin, B., Reiser, B. J., and Edelson, D. (2004). Scaffolding analysis: extending the scaffolding metaphor to learning artifacts. *J. Learn. Sci.* 13, 387–421. doi: 10.1207/s15327809jls1303_5

Singer, A., Montgomery, G., and Schmoll, S. (2020). How to foster the formation of stem identity: studying diversity in an authentic learning environment. *Int. J. STEM Educ.* 7, 57. doi: 10.1186/s40594-020-00254-z

Steele, C. M., Spencer, S. J., and Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. *Adv. Exper. Soc. Psychol.* 34, 379–440. doi: 10.1016/S0065-2601(02)80009-0

Vance, E. A. (2021). Using team-based learning to teach data science. *J. Stat. Data Sci. Educ.* 29, 277–296. doi: 10.1080/26939169.2021.1971587

Vygotskii, L. S. (1978). *Mind in society*. London: Harvard University Press.

Walton, G. M., Logel, C., Peach, J. M., Spencer, S. J., and Zanna, M. P. (2015). Two brief interventions to mitigate a “chilly climate” transform women’s experience, relationships, and achievement in engineering. *J. Educ. Psychol.* 107, 468–485. doi: 10.1037/a0037461

Weigelt, P., Denelle, P., Brambach, F., and Kreft, H. (2021). BotanizeR: A flexible r package with shiny app to practice plant identification for online teaching and beyond. *Plants People Planet* 4, 122–127. doi: 10.1002/ppp3.10226

Weintrop, D. (2021). The role of block-based programming in computer science education. In: *Understanding Computing Education (Vol 1). Proceedings of the Raspberry Pi Foundation Research Seminar Series* (Virtual event: Raspberry Pi), 1–8. Available online at: https://www.raspberrypi.org/app/uploads/2022/10/The_Role_of_Block_Based_Programming.pdf

Wertsch, J. V. (1984). The zone of proximal development: some conceptual issues. *New Dir. Child Adolesc. Dev.* 1984, 7–18. doi: 10.1002/cd.23219842303

Wiley, D., Bliss, T. J., and McEwen, M. (2013). Open educational resources: a review of the literature. In *Handbook of Research on Educational Communications and Technology* (New York, NY: Springer), 781–789.

Wiley, D., and Hilton, J. L. III. (2018). Defining oer-enabled pedagogy. *Int. Rev. Res. Open Distrib. Learn.* 19, 133–147. doi: 10.19173/irrodl.v19i4.3601

Yeager, D. S., Carroll, J. M., Buontempo, J., Cimpian, A., Woody, S., Crosnoe, R., et al. (2022). Teacher mindsets help explain where a growth-mindset intervention does and doesn’t work. *Psychol. Sci.* 33, 18–32. doi: 10.1177/09567976211028984

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 25 July 2023
DOI 10.3389/feduc.2023.1233337

Check for updates

OPEN ACCESS

EDITED BY
Wang-Kin Chiu,
The Hong Kong Polytechnic University, China

REVIEWED BY
Marisa Correia,
Polytechnic Institute of Santarém, Portugal
Hyuksoo Kwon,
Kongju National University, Republic of Korea

*CORRESPONDENCE
Monique Meier
✉ monique.meier@tu-dresden.de

†These authors have contributed equally to this work

RECEIVED 01 June 2023
ACCEPTED 11 July 2023
PUBLISHED 25 July 2023

CITATION
Thyssen C and Meier M (2023) 3D Printing as an element of teaching—perceptions and perspectives of teachers at German schools.
Front. Educ. 8:1233337.
doi: 10.3389/feduc.2023.1233337

3D Printing as an element of teaching—perceptions and perspectives of teachers at German schools

Christoph Thyssen[1†] and Monique Meier[2*†]

[1]Faculty of Biology, Department of Biology Education, University of Kaiserslautern-Landau (RPTU), Kaiserslautern, Germany, [2]Faculty of Biology, Department of Biology Education, Technische Universität Dresden, Dresden, Germany

Digital technologies that are very close to the teacher's analog field of activity, such as digital presentation, are increasingly taking place in the classroom, while digital, innovative technologies (e.g., 3D Printing) lacking such equivalents are used much less. Although such technologies are associated with more intense methodological and didactic changes, little is known about the extent to which 3D Printing is being used in German schools and how it is changing teaching and perspectives, which complicates the design of education and training measures. The use of such innovative technologies in the classroom is decisively influenced by the openness and acceptance of the teacher toward student-centered forms of learning and these technologies. The aim of the present study was to find out what expectations teachers (already) have about the use and potential of 3D Printing in the classroom and to what extent these are related to personal and/or external factors (e.g., 3D printers available in the school, number of STEM subjects). Therefore, an online-based questionnaire study was conducted with teachers in Germany ($N = 100$) who had different experiences with 3D Printing. The evaluation is based on descriptive, inferential and correlative analyses. Almost half of the teachers are equipped with 3D printers in their schools, while their use is even less widespread. In the perceptions of 3D Printing in the classroom from a methodological and didactic perspective, among other things, differences were revealed between teachers with different expertise in the knowledge and use of 3D Printing. In particular, the use of 3D Printing technology in their own lessons leads to a broader conception, especially with regard to the promotion of competencies. The results suggest theoretical models describing how to integrate 3D Printing into the classroom and concepts for 3D Printing trainings.

KEYWORDS

3D Printing, perceptions, STEM teacher, innovative technologies, teacher training

1. Introduction

How innovative is 3D Printing technology? A question that leads to different or differentiated answers depending on the perspective or area of application. In industry, 3D Printing technology is already present in many areas. Here, some have been and will continue to be more driven by developments in this technology than others. The branches of nutrition/food and fashion, as well as healthcare and aerospace, stand out in the application and research of new processes (Isi and Gurley, 2023). The latter are closely linked to areas of research in medicine. In particular, 3D Printing technology is a proven and established tool in medicine (Kalaskar, 2022) as well as in related areas of the life sciences. For example, biomedical materials that play an important role in organ transplantation, among others, are shaped using 3D Printing technology (Yan et al., 2018). 3D bioprinting is a promising approach for the production of complex biological constructs in biomedicine (Munaz et al., 2016). Another application example is 3D-printed biocarriers that help improve the efficiency of wastewater treatment (Sfetsas et al., 2021). There are many other examples of the use of 3D Printing technology in both scientific research and in industry. In line with ongoing technological developments, the use of 3D Printing technology still faces many challenges, especially in biology and medicine (e.g., in the development of printable materials, Zhou et al., 2020). In this respect, the potential of this technology is still unexploited even in the scientific field, despite its wide range of applications. For the related field of vocational education, it is obvious to include 3D Printing technology as a learning content and as a teaching-learning tool. In medical education, the use of 3D printed models in teaching human anatomy is well known, as are the associated effectiveness studies (Ye et al., 2020; Barreto et al., 2022; Ye et al., 2023). Acceptance of incorporating this technology into education is also high among students, while student knowledge of the use of 3D Printing technology for medical applications is very low (Wilk et al., 2020). Even with the innovation of 3D Printing technology in the medical and biological application field, there is a discrepancy in the integration of this technology in education. This also applies to school education, where not only content on the application of 3D Printing technology in the relevant subjects, but also the use of 3D Printing as a learning tool in the classroom appears to be less pronounced. There is a lack of well-founded data on this, both from a country-specific (e.g., Aslan and Celik, 2020) and an international perspective. In contrast to the continuously growing number of conceptual papers on the integration of 3D printers (e.g., Augusto et al., 2016; Monkovic et al., 2022; Oss Boll et al., 2023), the existing research on the use of 3D Printing technology in formal and informal education (e.g., Ford and Minshall, 2019) and the studies on the learning effectiveness of 3D Printing (e.g., Novak et al., 2021), it is only possible to make very limited statements about the actual use of this technology in the classroom and about teachers' ideas and attitudes toward its use. However, teachers' beliefs about technology have been identified as a key factor in the successful implementation of new or innovative technologies in the classroom (Sugar et al., 2004; Hew and Brush, 2007). With previous presentation and/or training on 3D Printing technology, teachers' attitudes toward the technology are preferentially positive, i.e., they would use the technology in their own classrooms; they address possible positive effects of using 3D technology in learning environments and/or recognize the potential of integrating this technology to transform open learning structures (Schelly et al., 2015; Yıldırım, 2018). The present study is based on teachers' perceptions of 3D Printing and its methodological and content-related integration into their subject teaching in lessons, but without any prior influence on these perceptions. The aim of this study is to gain insight into and describe the initial situation of teachers with regard to the integration of 3D Printing technology in the classroom, in order to derive recommendations for university teacher training and in-service training, as well as research opportunities. There is no comparable study for Germany, and the current status of the use of 3D Printing in German schools can only be inferred from published practical examples (e.g., Renner and Griesbeck, 2020; Bonorden and Papenbrock, 2022).

1.1. Outcomes of 3D Printing in education

From the perspective of 3D Printing experts, the integration of 3D Printing into teaching concepts requires competencies in the teacher and user especially in the area of 3D modeling and problem-solving competencies, creativity, and the knowledge of manufacturing and 3D Printing materials (Assante et al., 2020). Teachers who use 3D Printing technology in their classrooms also cite 3D modeling as an essential competence that students learn when participating in 3D projects, closely followed by the fostering of creative thinking and problem-solving skills, as well as technology skills (Trust and Maloy, 2017). According to the multiple competencies addressed, the active use of 3D Printing technology by students is very demanding in application of skills to implement a creative thinking and construction process—starting from an idea/problem, through modeling a solution, to printing a 3D object. With this technological complexity, the subject-based curricular learning as a result of the design/making process or in the use of the 3D objects does not take a back seat. Rather, a technology-based linking of science, technology, engineering, and mathematics (STEM) enables teaching and learning with 3D Printing in multidisciplinary, situated, subject-specific learning contexts (Pearson and Dubé, 2022). 3D Printing technology in school education is particularly closely associated with the promotion of STEM education (Ford and Minshall, 2019). In science, for example, this technology enables the understanding of complex systems, interactions and/or structures (e.g., in cell biology: Bagley and Galpin, 2015; e.g., in ecology: Kwon et al., 2020; e.g., in chemistry: Pernaa and Wiedmer, 2020); in mathematics, e.g., development of spatial visualization skills (Medina Herrera et al., 2019); in engineering, it is practical skills in the creation process (Chen and Cheng, 2021); and in engineering education, e.g., an engagement with sustainability and 3D Printing (To et al., 2023). While the expectations for 3D Printing technology are comparatively high, the research field for effective integration of 3D Printing technology into the curriculum is still very limited (Chen et al., 2023). Regardless of the discipline investigating the impact of 3D Printing technology on student learning, the learning potential of this technology is evident in STEM education as well as for non-STEM disciplines (Novak et al., 2021). In addition to creativity (e.g., Chien and Chu, 2018), spatial imagination (e.g., Wang et al., 2021), technical skills (e.g., Kwon, 2017), problem-solving skills and their linkage to creative thinking processes (e.g., Bicer et al., 2017), and cross- and interdisciplinary knowledge (Novak et al., 2021), communication and collaboration skills (especially in teaching visually impaired learners: e.g., Pantazis and Priavolou, 2017), motivation in learning (Kwon, 2017), and self-regulatory learning are also fostered in the implementation of 3D Printing projects. The latter is essentially accompanied by a mostly constructivist and hands-on as well as critically reflective and situated methodology in teaching and learning with 3D Printing (Pearson and Dubé, 2022). 3D Printing therefore has potential from both a methodological and a didactic point of view, and it is important to gain a fairly accurate insight into teachers' beliefs in these areas.

1.2. Integration 3D Printing in teaching

The 3D printer as a tool for printing prototypes and for the 3-dimensional visualization of ideas has experienced a strong boost in its integration into teaching as a result of the maker movement. Digital technologies and additive manufacturing processes, such as those used in 3D Printing, are characteristic of a wide range of making projects (Martin, 2015). Consequently, learning with 3D printers is closely related to maker-centered learning, which engages students in creative design processes in STEM disciplines (Hsu et al., 2017). Although the main goal in making is to produce a "product" that can be used, interacted with, or demonstrated (Martin, 2015), the process of designing, building, and producing is equally central to the active and problem-based learning emphasized by the maker movement (Martinez and Stager, 2013). A methodical integration of 3D Printing technology addresses both perspectives. On the one hand, it integrates the 3D object as a learning medium, i.e., it is available to teachers as a presentation medium and to learners for knowledge acquisition (Chen et al., 2023). In the learning process, the 3D printed object can be used as a subject-specific model, tool, spare part, visual/structural model or functional model, depending on the intended learning function and didactic-methodological integration (Meier et al., 2022). On the other hand, learners can be enabled to design and produce their own 3D printed objects. The design and production process spans between the 3D printer as a device and tool of the subject sciences and the 3D object as a product (e.g., the material model of an original; Meier et al., 2022). Embedded in a subject-specific context, learners go through a technology-supported model-building process when designing and printing: starting with the original, via an idea and a mental model, to the virtual model and the context-related application of a printed 3D object (with a biology example: Meier and Thyssen, 2021). If (also) the design process up to printing comes more into focus, this requires not only technological competencies on the part of the teachers, but also a (partial) "opening" of the traditional, teacher-centered teaching structures. In the synopsis of studies on learning by means of 3D Printing technology, problem- and project-based learning are mentioned as teaching concepts (Novak et al., 2021), as well as a constructivist and design/making-oriented understanding of learning (e.g., design thinking: Greenhalgh, 2016). While a problem-based learning approach is not necessarily linked to the production of 3D printed objects by learners, it can guide engagement with a subject content when combined with project-based learning. Project-based learning with integrated 3D Printing technology, on the other hand, is directly linked to learner engagement in the creative, communicative, and iterative process of producing a 3D printed object (Novak et al., 2021). Together with the integration of problem-based learning, among other things, this creates opportunities for inquiry-based learning in which learners solve real-world problems that span multiple disciplines (Ali et al., 2019). What stands out in the design process is the active (co-)design participation of the students. On the part of the teachers, this makes it necessary to plan and create various, individually adapted support activities/strategies. These include not only facilitating the use of technology, but also supporting collaboration and communication, design, and the understanding of the subject matter (Chen et al., 2023). Against this background, it is essential to know the teachers' perspective on the possible methodological and didactic integration of 3D Printing and their assessment of the possibilities of developing students' competencies in the above-mentioned areas, which requires appropriate data collection and analysis.

2. Research questions and hypothesis

Technical equipment often plays a central role or is a major obstacle for teachers when dealing with the integration of digital technology in subject lessons (in addition to a lack of competence and confidence; Bingimlas, 2009). Without access to the technology, an examination of it seems obsolete—a circumstance that does not apply equally to every technological approach. 3D Printing can also be a tool in the learning process without the physical device, for example by emphasizing 3D modeling and/or outsourcing printing to external service providers (Kantaros et al., 2022). However, assuming that an available device triggers and influences teachers' planning and thinking processes for teaching with 3D Printing, its occurrence in schools would be a first starting point for further studies or training. For Germany, there is a lack of knowledge regarding the country-specific expression of this initial technological condition as well as the associated interest in the use of 3D Printing and further training. Therefore, this study exploratively addresses the following research questions:

> *Q1a*: To what extent is 3D Printing technology a part of the digital equipment in schools, and is its use by teachers a widespread practice in German schools?
>
> *Q1b*: How is teacher interest in 3D Printing measured in terms of a desire for 3D Printing equipment in their own school and the willingness or rather participation in educational training?

A number of literature reviews have described the learning effects associated with the use of 3D Printing (see section 1.1). Increased motivation and creativity toward the 3D object and the design process (e.g., Bécar et al., 2017) as well as the promotion of subject-specific competences through the integration of 3D Printing in the classroom are examples of (presumed) effects (Ford and Minshall, 2019). However, the question remains open as to what potential teachers see in the use of 3D Printing and how their perceptions in this area are influenced by individual parameters. The following research question and hypotheses are posed:

> *Q2*: Are there differences in teachers' perceptions of the benefits of 3D Printing for student competence development according to their age (1), the subjects they teach (2) and/or 3D Printing experience/expertise (3)?
>
> *H2.1*: *Younger teachers do not differ from older teachers in their perceptions of the competencies that 3D Printing fosters in students.* [In line with the lack of empirical evidence on the relationship between age and, for example, perceptions of information and communications technology (ICT) related competencies (e.g., Guo et al., 2008)].
>
> *H2.2*: *Teachers who teach at least one or more STEM subjects differ from teachers who do not teach STEM subjects in their perceptions of the competencies that 3D Printing fosters in students*

(Corresponding to the different use of subject-specific digital media by teachers who teach a STEM subject and those who do not (Lorenz and Eickelmann, 2022), the technology for 3D Printing is also based on a subject-specific STEM orientation, which may lead to differences in teachers' perceptions of competence).

H2.3: *Teachers with more experience/expertise with 3D Printing in an educational context differ in their perceptions of the competencies that 3D Printing fosters in students* (According to Trust and Maloy (2017), in the present study these are particularly creativity, problem solving, and technological literacy).

With the integration of 3D Printing into teaching, the design process up to printing, the printed object or even the printer itself becomes the focus of the subject-related learning process (e.g., Pearson and Dubé, 2022; see section 1.2). The extent to which these integration/learning scenarios for 3D Printing are known or perceived by teachers, and the possibilities they see in the methodological and didactic design, can only be guessed at. The empirical field is based on subject-specific studies in which teachers are explicitly exposed to 3D Printing technology before they are asked to execute their ideas or own teaching scenarios for integrating 3D Printing (e.g., Trust and Maloy, 2017; Novak and Wisdom, 2020). For our study, the teachers' perceptions, without the influence of a 3D-supported learning environment or in-service training, are explored descriptively investigated with the following research question:

Q3: What are teachers' perceptions of the possibilities of integrating 3D Printing into the classroom from a didactic and/or methodological perspective?

3. Materials and methods

3.1. Instrument and data collection

Data collection was online-based and anonymous using questionnaires in 2022. The surveys were sent to and distributed at schools preferably located in the local area of the researchers in Germany. The schools contacted were randomly and equally selected with respect to a presumed inventory of 3D printers, derived from the respective information on the school's homepage. The aim was to generate a heterogeneous sample in terms of 3D Printing expertise, age and subject (see section 3.2). In this respect, there were no restrictions on participation in the survey.

The first block of the questionnaire for socio-demographic data (e.g., gender, age, school type, professional duration) is followed by eight sections of questions directly related to 3D Printing technology, with a total of 50 items (Table 1). The item format consists of content-based choice responses (single-choice or multiple-choice) as well as open and closed formats, the latter with an 8-point Likert scale (for interest: from 1 = no interest to 8 = very high interest, for consent: from 1 = strongly disagree to 8 = strongly agree, for knowledge: from 1 = not at all expressed to 8 = very highly expressed) and the option for no answer. On the one hand, the development of the questionnaire was theoretically and empirically driven, especially with regard to the items on competence promotion (see section 1.1) and didactic-methodological integration (see section 1.2). Several competency domains promoted by 3D Printing were derived from empirical studies/results for the survey, such as creativity, conceptual understanding, problem solving, and motivation [e.g., Chen and Cheng, 2021; (D) in Table 1]. The items on the methodological-didactic integration of 3D Printing in the classroom [(E)–(G) in Table 1] are based on the theoretical multidimensional concept of Meier et al. (2022) as well as empirical studies in this area (e.g., Novak et al., 2021). The integration of a questionnaire block on workshop participation and expectations for further education on 3D Printing ties in with a need that is not met or should receive more attention in the German as well as international field (e.g., Choi and Kim, 2018; Diepolder et al., 2021; (C) in Table 1). In addition to the theoretical connection, the construction of the questionnaire was based on the exchange and consensus of a multidisciplinary working group consisting of researchers with expertise in 3D Printing from four German universities and three scientific disciplines (Biology, Chemistry, and Physics). This group created an item pool from items related to the mentioned fields and feedback from teachers in the field. By selecting items, a questionnaire was created, adapted, and finalized with appropriate items in several cycles.

The question sections formed are generally not oriented to a strict direction in terms of content. The items on possible competence development and on the general perception of 3D Printing are intended to cover a wide field in order to be able to record different perceptions. Consequently, these do not form a unidimensional scale and no reliability analysis is performed. An evaluation is then done at the level of the individual item. In contrast, for the three specific areas of integration of 3D Printing technology (methodology, didactics, sustainable development) in teaching, both the individual items and their composition in a corresponding scale are analyzed. In the total sample (excluding missing statements), the scales consistently show satisfactory to good reliability, with a Cronbach's $\alpha > 0.70$ (Bühner, 2011): methodical integration of 3D Printing: $\alpha = 0.775$, $N = 93$; didactic integration of 3D Printing: $\alpha = 0.839$, $N = 82$; promoting sustainable development (SD): $\alpha = 0.907$, $N = 63$.

3.2. Sample

In total, 100 teachers (51% female, 47% male, 2% not specified) were asked about their opinions and perceptions of 3D Printing in the classroom as well as the status of digital equipment for 3D Printing at their respective schools (see section 3.2). The mean age of the participants is 44.3 years ($SD = 9.78$). The mean number of years of professional experience in the total sample is 12.48 years ($SD = 9.07$). Twenty seven teachers (27%) are employed at secondary schools ("Haupt-/Sekundarschule"), 35 teachers (35%) work at comprehensive schools ("Gesamtschule"), 32 teachers (32%) are employed at grammar schools ("Gymnasium") and 5 teachers work at other types of school (e.g., vocational school). According to the research questions (Q2) and hypotheses (H2.1–H2.3), the total group of teachers surveyed in this study was divided into subgroups.

3.2.1. Forming age groups (H2.1)

Three groups were created based on age: up to and including 40 years ($n = 34$, 31.2%), 41 to 50 years ($n = 36$, 33.0%), over 50 years ($n = 29$,

TABLE 1 Structure and design of the questionnaire.

Question section	Item count	Example items	Response format
(A) Expertise	4	My technological knowledge is…	Knowledge scale
		My informatics knowledge is…	
		My 3D Printing knowledge is …	
		Have you used a 3D printer yourself? (No/Yes, private!/Yes, for lesson planning!/Yes, in lesson!)	Multiple choice
(B) Equipment	3	Are there 3D printers at your school? (No/Yes, one!/Yes, several!/I do not know!)	Multiple choice
		Would you like to see a 3D printer purchased at your school? (No/Yes/Maybe/I cannot judge.)	Multiple choice
(C) In-service training	3	Assess your interest in attending a training seminar on the use of 3D Printing in the classroom.	Interest scale
		For training seminars on the use of 3D Printing in the classroom, here's what I'd like to see…	Open
(D) Competence promotion	10	*The following competencies can be particularly promoted in learners with the use of 3D Printing in the classroom:* e.g., creativity, model competence, problem solving competence	Consent scale
(E) General about 3D Printing	9	*When I hear the "3D Printing" term, I think…*	Consent scale
		…to the physical device.	
		…to the printed product.	
		…to the design process that can be integrated.	
		…rather to a field for other subjects.	
	9	*When thinking about 3D Printing, I see possibilities…"* e.g.,	Consent scale
		The methodical integration into the classroom.	
		The content/didactic integration into the classroom.	
		The creation of individualized/ differentiated approaches to learning	
		With 3D Printing for and in the classroom, I associate…	Open
(F) Methodical integration of 3D Printing	3	*When thinking about 3D Printing, I see opportunities for methodological integration…*	Consent scale
		Production of 3D models and 3D objects.	
		Production of experimental material.	
		To involve students in activities related to 3D Printing.	
(G) Content-related/didactic integration of 3D Printing	4	*When thinking about 3D Printing, I see possibilities for contextual and therefore didactic integration…*	Consent scale
		Technology in the disciplines corresponding to the subjects I teach.	
		Everyday context.	
		Context of societal changes and challenges.	
		Sustainability context.	
(H) Promoting sustainable development (SD)	5	*In thinking about 3D Printing, I see opportunities to promote sustainable development, through…*	Consent scale
		Production on site.	
		Printing of individual spare parts.	
		Production of parts for upcycling constructions.	
		Recycling of plastics for printing polymers.	
		SD concepts on 3D Printing.	

26.6%). Based on an assumed average age of 28 to 30 years at the beginning of the teaching profession (after completion of the practical phase and 2nd state examination), the interval size is at least 10 years.

3.2.2. Forming subject groups

Additionally, based on the data describing the subjects taught by each in-service teacher, three groups were created: no STEM subject

($n = 23$), one STEM subject ($n = 38$) and at least two STEM subjects ($n = 39$).

3.2.3. Forming experience levels (H2.3)

Furthermore, teachers were divided into groups according to their self-reported technological, informational and 3D Printing knowledge [see (A) in Table 1]. Teachers with data above the mean of the 3D Printing knowledge scale (M = 4.5) formed the 3D Printing Expert group, and teachers with data below the mean formed the 3D Printing Novice group (Table 2). Based on the self-reported knowledge, the subgroups will always be referred to as 3D Printing Experts and 3D Printing Novice in the further course of the article. To examine correlations of 3D Printing expertise, experience with 3D Printing in the classroom was used as a grouping variable in addition to self-reported knowledge in this area. All teachers who reported having used a 3D printer in their own classrooms (3D Users: $n = 22$) were compared with other teachers who had no classroom experience with 3D Printing (3D Non-Users: $n = 78$). Throughout the rest of the article, these will be referred to as 3D User and 3D Non-User. When comparing the size of the resulting groups for the three different knowledge domains (Table 2) the data indicates that the fraction of Novices is increasing from Technology via Computer Science to 3D Printing. As a result, the amount of knowledge and the number of contact persons for 3D Printing in the teaching staff is the lowest. Novices in 3D Printing also have the lowest level of knowledge. There are positive correlations (Spearman-Rho) between the three types of knowledge with strong effects in all cases (r_s between 0.652 and 0.711, all $p < 0.001$). For example, low levels of technology knowledge are associated with low levels of 3D Printing knowledge. Based on self-reported knowledge and usage there are significant differences between all three knowledge domains in both 3D Printing expertise subsamples, namely Users/Non-Users and Experts/Novices. When comparing 3D Users with 3D Non-Users the groups differ significantly (Technology Knowledge with $U = 373.5$, $z = -4.08$, $r = 0.41$; Computer Science Knowledge with $U = 346.5$, $z = -4.30$, $r = 0.43$; 3D Printing Knowledge with $U = 36.5$, $z = -7.19$, $r = 0.71$, all $p < 0.001$). In this respect, the 3D Non-Users consistently rate their skills lower than the 3D Users. The same ist true for the 3D Printing Novices and 3D Printing Experts (Technology knowledge with $U = 282.0$ $z = -5.53$, $r = 0.55$; Computer Science Knowledge with $U = 271.0$, $z = -5.61$, $r = 0.56$; 3D Printing Knowledge with $U = 0$ (due to group definition), $z = -8.05$, $r = 0.81$, all $p < 0.001$).

3.2.4. Knowledge and subjects (H2.2 & H2.3)

In addition, there are significant differences in the knowledge groups (Table 2) according to the number of STEM subjects they taught. The group with teachers who teach two STEM subjects rate their knowledge in all three areas significantly higher than teachers in the other two groups [Technology: $H(2) = 16.77$, $p < 0.001$; Computer Science: $H(2) = 16.10$, $p < 0.001$; 3D Printing: $H(2) = 17.03$, $p < 0.001$]. *Post hoc* Tests (Dunn-Bonferroni-Tests) show significant differences to both groups with moderate effects (Technology: 2 STEM vs. no STEM $z = -3.66$, $p < 0.001$, $r = 0.47$ and vs.1 STEM $z = -3.22$, $p = 0.001$, $r = 0.37$; Computer Sciences: 2 STEM vs. no STEM $z = -3.60$, $p < 0.001$, $r = 0.46$ and vs.1 STEM $z = -3.14$, $p = 0.002$, $r = 0.36$; 3D Printing: vs. no STEM $z = -3.63$, $r = 0.46$ and vs.1 STEM $z = -3.33$, $r = 0.38$ for both $p < 0.001$).

There are no significant differences between the three age groups in self-reported knowledge of technology, computing and 3D Printing.

3.3. Data analysis

Descriptive analyses are used to quantitatively describe the baseline situation in terms of 3D Printing equipment and teachers' perceptions. Frequencies and location and dispersion parameters [median (Mdn), mean (M), and standard deviation (SD)] will be reported. Inferential statistical procedures are used to test for group differences and correlations. Since almost all data were not normally distributed (Shapiro-Wilks and Kolmogorov–Smirnov tests, $p < 0.05$), non-parametric procedures were used to compare groups. Depending on the number of independent groups, e.g., 3D Users vs. 3D Non-Users, the Mann–Whitney *U*-test or, in the case of more than three independent groups, the Kruskal-Wallis test was used. *Post-hoc* tests (Dunn-Bonferroni test) are used to specify group differences. In this case, the adjusted value of p is quoted. The Wilcoxon test is used to analyze differences in the overall sample, e.g., for comparing the different perceptions of the didactic and methodological integration of 3D Printing. Spearman rank correlation (r_s) was used to test correlations. Significance level was set to $p \leq 0.05$. The effect sizes are evaluated according to Cohen (1992).

4. Results

4.1. Availability and usage of 3D Printing (Q1a)

Regarding the availability of 3D printers [see (B) in Table 1], a disproportion between the types of schools can be observed. Teachers at grammar schools ("Gymnasium": 66%) and comprehensive schools ("Gesamtschule": 54%) in particular indicate that they have one or more printers. Only one secondary school ("Haupt-/Sekundarschule") teacher states that there are 3D printers at the school. Relative to the

TABLE 2 Expert and novice subgroups related to technological, informational and 3D Printing knowledge.

Groups	Technology knowledge			Computer science knowledge			3D Printing knowledge		
	N	*Mdn*	*M ± SD*	*N*	*Mdn*	*M ± SD*	*N*	*Mdn*	*M ± SD*
Experts in…	62	6	6.31 ± 0.985	48	6	6.15 ± 1.052	27	6	6.41 ± 1.047
Novices in…	38	3	2.97 ± 1.052	52	3	2.79 ± 0.97	73	1	1.62 ± 0.922
3D Users	22	7	6.45 ± 1.405	22	6	6.05 ± 1.588	22	6.5	6.50 ± 1.185
3D Non-Users	78	5	4.64 ± 1.851	78	4	3.94 ± 1.819	78	1	1.90 ± 1.392

total sample, 45% of the teacher's report having at least one 3D printer in school (20% of teachers report having more than one), while 49% of the teachers do not have a printer in school (the rest is unsure). Just 22% of the teachers already have used 3D Printing for teaching purposes, either in preparing lessons or during lessons itself, while 22.9% already have used 3D printers for private purposes. Less than 2% of the teachers who have already used 3D Printing for lesson related purposes aren't STEM-teachers. Just 2 teachers belonging to the 3D Printing Novice group stated that they already integrated 3D Printing into their lessons, while 20 of the 3D Printing Experts did. In terms of the presence or absence of one or more 3D printers in the schools, a significant correlation can be found with classroom use, $r = 0\ 0.527$, $p < 0.001$, $N = 100$ (with a strong effect), as well as with 3D Printing Expert/Novice knowledge, $r = 0.394$, $p < 0.001$, $N = 100$ (with a moderate effect). In particular, the 3D Printing Experts state that they have one (26%) or more 3D printers (60%) available in school. While a larger group of 3D Printing Novices do not have a 3D printer (61%), 25% say they have one and 5% say they have several 3D printers in school.

4.2. Interest for 3D printers and in-service training for 3D Printing (Q1b)

Among teachers who do not have a 3D printer in their school, 44% would like to purchase one. Of those who already have one or more 3D printers in their school, 71% would like to purchase another 3D printer. In total, 56% would like to purchase (another) 3D printer, 23% say that such a purchase might be necessary, and only 8% (25% of these teachers already have a 3D printer at school) do not want to purchase one. The 3D Printing Novices ($N = 73$, $Mdn = 2$, $M = 2.62$, $SD = 0.91$) would rather appreciate a purchase than the 3D Printing Experts ($N = 27$, $Mdn = 2$, $M = 2.04$, $SD = 0.52$), U = 613.5; $z = -3.175$, $p < 0.001$, $r = 0.32$.

Only 14% of the teachers (based on responses from 100 respondents) have participated in 3D Printing in-service trainings [see (C) in Table 1]. The reasons given by the remaining teachers for not attending such training events were (still) a lack of interest (34%) and a lack of suitable offers (22%). Likewise, the limited time available to pursue such training plays an important role for teachers (12%). Interest in further training on 3D Printing is fairly evenly distributed among the group of respondents ($N = 74$), with 55.4% indicating no to little interest and 44.6% indicating high to very high interest. The group of 3D Printing Experts shows significantly higher interest in trainings ($Mdn = 5.5$, $M = 5.15$, $SD = 2{,}22$) than the group of 3D Printing Novices ($Mdn = 3$, $M = 3.67$, $SD = 2{,}12$) in 3D Printing (U = 390.5; $z = -2.676$, $p = 0.007$, $r = 0.311$). There are no differences in interest in 3D Printing education among the groups divided by age and number of STEM subjects.

4.3. Perceptions on putative competence development by 3D Printing (Q2)

For all competency areas surveyed regarding their ability to promote them with 3D Printing (see (D) in Table 1), the mean scores across all teachers were above the scale mean ($4.66 < M < 6.61$, Table 3). In the perception of the teachers, the use of 3D Printing in the classroom is mainly beneficial for the development of general technical skills and competencies in modeling. On the other hand, there is a lower value in the competence areas of communication and cooperation, which could benefit from the integration of 3D Printing in the classroom (Table 3).

4.3.1. Age and Subjects (H2.1 & H2.2)

When comparing groups of teachers of different age and number of STEM subjects, no significant differences were found between the groups in terms of their rated potential for promoting competencies. This is also the case when grouped by technological or computer science knowledge.

4.3.2. Experience levels (H2.3)

For the pairs formed with different levels of expertise, 3D Printing Novices/Experts and 3D Non-Users/Users, some differences were found in the reported scores of the competencies that can be developed through 3D Printing, with the more experienced group rating the development possibilities higher (Table 3). Based on the reported 3D Printing Knowledge significant differences can be found for promoting competencies in the areas of creativity (U = 702.500, $z = -2.041$, $p = 0.041$, $r = 0.21$), scientific inquiry (U = 307.000, $z = -3.488$, $p < 0.001$, $r = 0.41$), problem solving (U = 358.500, z = −4.561, $p < 0.001$, $r = 0.48$), and general digital competencies (U = 670.000, $z = -2.183$, $p = 0.029$, $r = 0.22$). The reported 3D Printing Knowledge correlates only with perceptions on fostering scientific inquiry with a medium effect ($r = 0.357$, $p = 0.002$) and problem-solving competencies ($r = 0.383$ $p < 0.001$). If the classification is based on the integration of 3D Printing into lessons, significant differences can also be found for promoting competencies in the area of creativity (U = 554.500, $z = -2.437$, $p = 0.015$, $r = 0.25$), scientific inquiry (U = 286.500, Z = −3.293, p < 0.001, r = 0.39), problem solving (U = 243.000, $z = -4.944$, $p < 0.001$, $r = 0.52$) and general digital competencies (U = 494.000, z = −2.857, $p = 0.004$, r = 0.29). In addition, significant differences are shown in the perception of promoting social (U = 535.000, $z = -2.177$, $p = 0.030$, $r = 0.22$) and communication competencies (U = 562.000, $z = -2.072$, $p = 0.038$, $r = 0.21$).

4.4. Perceptions about the methodical and didactical integration of 3D Printing in the classroom (Q3)

With regard to the perception of the possibilities of using 3D Printing [also with students, see (E) in Table 1], the methodological integration for the production of 3D models and objects ($Mdn = 7$, $M = 6.75$, $SD = 1.77$) is at the top of the list, while the perspective of using it as a teacher for lesson planning without involving students, for example, plays a lesser role ($Mdn = 2$, $M = 3.00$, $SD = 2.02$). When asked for a general assessment with a single item, teachers were very similar in their perceptions of the potential for integrating 3D Printing methodologically ($Mdn = 6$, $M = 5.38$, $SD = 2.34$) and didactically ($Mdn = 6$, $M = 5.28$, $SD = 2.22$). In contrast to this finding the values obtained by using the scales for methodological integration [see (F) in Table 1: $Mdn = 6.66$, $M = 6.34$, $SD = 1.57$] and didactic integration [see (G) in Table 1: $Mdn = 4.75$, $M = 4.68$, $SD = 1.74$] differed significantly with a strong effect [Wilcoxon, $z = -7.88$,

TABLE 3 Perceptions of competency areas that 3D Printing can foster as reported by panel (total sample), 3D Printing Novice/Expert, and 3D Non-User/User (pairs of values in bold indicate significant differences).

Competency areas	Total sample				3D Printing Novice				3D Printing Experts				3D Non-User				3D User			
	Mdn	*M*	*SD*	*N*	*Mdn*	*M*	*SD*	*N*	*Mdn*	*M*	*SD*	*N*	*Mdn*	*M*	*SD*	*N*	*Mdn*	*M*	*SD*	*N*
Creativity	7	6.61	1.64	97	7	**6.36**	1.79	70	7	**7.26**	0.90	27	7	**6.37**	1.75	75	8	**7.41**	0.80	22
Scientific inquiry	7	6.15	1.73	72	6	**5.63**	1.78	46	8	**7.08**	1.20	26	6	**5.72**	1.78	50	8	**7.14**	1.13	22
Problem solving	7	6.23	1.77	92	6	**5.74**	1.82	65	8	**7.41**	0.84	27	6	**5.79**	1.78	70	8	**7.64**	0.58	22
Cooperation/social	5	5.05	1.96	92	5	4.86	1.96	65	5	5.52	1.93	27	5	**4.79**	1.96	70	6	**5.91**	1.74	22
Modeling	7	6.52	1.44	95	7	6.37	1.50	68	7	6.89	1.25	27	7	6.37	1.50	73	7	7.00	1.15	22
Subject knowledge	5	5.38	1.70	95	5	5.19	1.69	68	6	5.85	1.66	27	5	5.23	1.74	73	6	5.86	1.46	22
Communication	5	4.66	2.12	94	4	4.45	2.18	67	5	5.19	1.90	27	4	**4.40**	2.15	72	6	**5.50**	1.82	22
SD	5	5.10	1.96	71	5	5.00	2.05	53	6	5.39	1.69	18	5	4.95	2.01	56	6	5.67	1.68	15
General digital	6	6.27	1.44	96	6	**6.07**	1.46	69	7	**6.78**	1.28	27	6	**6.05**	1.44	74	7	**7.00**	1.20	22

$p < 0.001$, $r = 0.81$, $n = 94$]. While the values derived from the two scales increased for methodological integration, they decreased for didactic integration. When comparing 3D Printing Novices and Experts, the perspectives on methodological and didactic integration are different, regardless of which of the two indicators (scale or single item) is analyzed (Table 4).

Perceptions of methodological and didactic integration correlate significantly with each other on an individual item basis with strong effect $r = 0.881$, on a scale level only with $r = 0.675$ (all $p < 0.001$). In some options for methodological and didactic integration, the ratings of 3D Printing Experts differ significantly from those of the Novices (Table 5). Looking at the perspective on a 3D printer in terms of associated thoughts, there are significant differences for 2 items. For the other items, the means and medians of the ratings of what teachers think when they hear the term "3D printing" are in a range between $3 < Mdn < 8$ and $3.68 < M < 7.07$ with minima for the items different printing processes and field for other colleagues, while maxima were observed for the items physical device and printed product ($Mdn = 8$, $M = 6.69$, $SD = 1.88$ for physical device, $Mdn = 7.5$, $M = 7.07$, $SD = 1.29$ for printed product). The results for thoughts associated with the 3D design process were $Mdn = 5$, $M = 4.94$, and $SD = 2.438$.

With regard to the perception of the possibilities for integrating 3D Printing into teaching, i.e., from a methodological, didactic and/or sustainable development (SD) perspective, 3D Printing Experts rate 3D Printing significantly differently on 7 items (Table 5). For these items, the expert ratings were higher than the novice ratings, both in terms of median and mean.

When testing for groups of teachers who had or had not used 3D Printing in the classroom, in addition to the same items with significant differences specifically related to knowledge of 3D Printing, one additional item shows significant differences (Table 5B), related to the sustainable production of spare parts. In this case, 3D Printing Users rated higher. There are no differences between the expertise groups for the other 4 ways/items in which 3D Printing can be used to promote sustainable development. The scores are between $4 < Mdn < 6$ and $4.41 < M < 5.59$ for the items covering printing spare parts, upcycling constructions, recycling and SD concepts. The scoring for printing spare parts ($Mdn = 7$, $M = 6.18$, $SD = 2.074$) and printing material for experiments as an item covering methodological aspects ($Mdn = 6.5$, $M = 6.15$, $SD = 1.835$) do not differ when compared using Wilcoxon test.

4.5. Perceptions of groups teaching different numbers of STEM subjects

Data on the use of 3D printers show a tendency that higher use, particularly in the classroom, is observed when two or more STEM subjects are taught. The comparison of 3D Printing Users and Non-Users shows an identical number of users a similar distribution in the private use of 3D Printing in contrast in addition to the difference in educational use (Table 6).

Furthermore, there are clear differences in whether teachers see 3D Printing as a domain of their own or other subjects [see (G) in Table 1] when no or at least 2 STEM subjects are taught (Table 7). According to the lowest *Mdn* values teachers with 2 STEM subjects, 3D Printing is more likely to be seen in the STEM subjects. In line with this, STEM teachers are also much more likely to classify 3D

TABLE 4 Perceptions of groups with different levels 3D Printing knowledge on methodological and didactic integration of 3D Printing into the classroom.

Integration—item or scale	3D Printing Novice				3D Printing Expert				Testing statistics			
	Mdn	*M*	*SD*	*N*	*Mdn*	*M*	*SD*	*N*	*U*	*z*	*p*	*r*
Methodical—one item [see (E) in Table 1]	6	4.97	2.366	61	7	6.30	2.09	27	549.5	−2.523	0.012	0.27
Didactical—one item [see (E) in Table 1]	5	4.60	2.199	55	7	6.73	1.46	26	317.0	−4.087	<0.001	0.45
Methodical—scale [see (F) in Table 1]	6.33	6.01	1.62	71	7.33	7.19	1.02	27	488.0	−3.767	<0.001	0.38
Didactical—scale [see (G) in Table 1]	4.25	4.27	1.67	68	6.00	5.75	1.47	26	437.0	−3.783	<0.001	0.39

TABLE 5 (A) Perceptions of groups with different levels of 3D Printing Expertise (based on self-reported knowledge) on the integration of 3D Printing in the classroom for items that show significant differences only for different levels of 3D Printing knowledge but not for different levels of technology or computer science knowledge; (B) Additional items that show significant differences only for different 3D User/Non-User in class but not for different 3D Printing, technology or computer science knowledge.

PART A	3D Printing Novice				3D Printing Expert				Testing statistics			
	Mdn	*M*	*SD*	*N*	*Mdn*	*M*	*SD*	*N*	*U*	*z*	*p*	*r*
When I hear the "3D Printing" term, I think… [see (E) in Table 1]												
…from an uninformed perspective	3	3.92	2.513	60	1	2.17	1.800	23	381.5	−3.205	0.001	0.35
…to a topic for sustainability	3	3.64	2.291	70	6	5.67	2.000	27	483.0	−3.752	<0.001	0.38
When thinking about 3D Printing, I see possibilities… [see (E) in Table 1]												
…to the linkage to curricular areas (none to many).	4	3.78	2.074	64	5	5.38	1.941	26	484.0	−3.129	0.002	0.33
…to the creation of individualized/ differentiated approaches to learning.	6	4.99	2.239	67	7	6.46	1.476	26	540.0	−2.870	0.004	0.30
…to the promotion of SD.	4	3.83	2.295	58	5	5.19	2.245	26	500.0	−2.486	0.013	0.27
When thinking about 3D Printing, I see opportunities for methodological integration… [see (F) in Table 1]												
…production of 3D models and 3D objects.	7	6.53	1.839	70	8	7.35	1.413	26	586.0	−2.838	0.005	0.29
…to involve students in activities related to 3D Printing.	6	5.58	2.199	69	8	7.44	0.934	27	427.0	−4.239	<0.001	0.43
When thinking about 3D Printing, I see possibilities for contextual and therefore didactic integration… [see (G) in Table 1]												
…everyday context.	4	4.08	2.010	66	6	5.73	1.909	26	470.0	−3.397	0.001	0.35
In thinking about 3D Printing, I see opportunities to promote sustainable development, through… [see (H) in Table 1]												
…production on site.	6	5.34	2.181	65	7	6.23	2.026	26	621.0	−1.998	0.046	0.21
PART B	**3D Non-User**				**3D User**				**Testing statistics**			
In thinking about 3D Printing, I see opportunities to promote sustainable development, through… [see (H) in Table 1]												
…printing of individual spare parts.	7	5.91	2.16	69	8	7.00	1.54	22	516.5	−2.318	0.020	0.24

TABLE 6 3D Printer usage of 3D Users/Non-Users and teachers teaching different numbers of STEM subjects [see (A) in Table 1].

N	0 STEM subject	1 STEM subject	At least 2 STEM subjects	3D User	3D Non-User
	23	38	39	22	78
None	17	28	18	0	63
Private	6	8	16	15	15
For lesson planning	2	3	11	14	2
In lessons	2	5	15	22	0

TABLE 7 Perceptions of groups teaching different numbers of STEM subjects.

3D as a...	0 STEM subject				1 STEM subject				At least 2 STEM subjects				Testing statistics					
	Mdn	*M*	*SD*	*N*	*Mdn*	*M*	*SD*	*N*	*Mdn*	*M*	*SD*	*N*	*H(2)*	*p*	*GP*	*z*	*p*	*r*
…field for other subjects.	5	4.7	2.548	23	4	4.03	2.284	38	2	3.08	2.186	38	7.310	0.026	**0 vs. 2**	2.569	0.028	0.33
…corresponding science technology in their subjects.	4	4.24	2.256	21	5	5.31	2.054	36	6	5.85	1.987	34	7.324	0.026	**0 vs. 2**	−2.706	0.02	0.36
…field for SD concepts.	3	3.42	2.364	19	5	5.22	2.063	27	5	4.53	2.091	19	6.961	0.031	**0 vs. 1**	−2.638	0.025	0.39

Printing as a corresponding science technology in their subject. With regard to a connection to SD, the picture is slightly different, as significant differences can only be observed between teachers without and with a single STEM subject. The latter see a stronger linkage.

5. Discussion and conclusion

Technological innovation is making its way into education, albeit slowly but steadily. How innovative 3D Printing technology is perceived from an educational and teaching perspective has been little studied. This study aims to provide some initial insights for Germany. Of particular interest are teachers' perceptions of the 3D printer as a teaching and learning tool in terms of skills development and the methodological and didactic integration of 3D Printing technology in the classroom. A description of the current status of 3D printers in German schools and their integration into subject lessons includes on the one hand on the equipment (Q1a) and on the other hand, of course, on the users of this technology (Q1b–Q3). Teachers are the driving force behind digitalization processes and efforts in schools. Their concepts and decisions to integrate digital technologies into the planning and delivery of teaching are influenced by many factors. These include attitudes toward digital technologies, as well as pedagogical knowledge and perceptions of effective integration in teaching, and their own technological skills (Ertmer et al., 2015).

5.1. Specifications of the sample in terms of 3D Printing expertise

In order to investigate the research questions and hypotheses (section 2), a heterogeneous sample of teachers is used, with varying numbers of STEM subjects and expertise, e.g., in technology, computer science and 3D Printing knowledge. In line with the study's focus on 3D Printing technology, the expertise of the teachers surveyed in this area is included in the analyses in the form of self-reported knowledge and use of 3D Printing in their own teaching. For this purpose, groups of Experts are compared with Novices and groups of Users with Non-Users. The decisive feature and legitimation for this grouping are the significant differences that exist in the self-reported areas of knowledge and use in teaching. Here, the scores of Experts and Users are consistently higher than those of Novices and Non-Users. Drossel et al. (2017) report that self-efficacy in preparing lessons involving the use of ICT is the only significant predictor of the use of computer use in schooling that is found in all countries surveyed. Like our data, their models also show no significant role for age, but experience in using ICT was one of the factors with the highest impact. The postulated differences between the groups in the context of 3D Printing may arise from the transformative, constructivist ideas attributed to Experts for designing digitally supported instruction in which they are consultative and open to new ideas (Berg et al., 1998; Meskill et al., 2002).

5.2. Current status on 3D Printing in German schools (Q1)

The 3D printer is no longer a newcomer either, and the equipment in German schools looks promising. About half of the teachers surveyed in this study said they had one or more 3D printers in their school. While there is room for improvement, especially in the much less well-equipped secondary schools ("Gymnasien"), this already opens up some possibilities for integrating this technology. In terms of both school type and level of use, the current picture in Germany is roughly supported by findings from other countries (Choi and Kim, 2018). Assuming that equipment has increased over the years, the main difference with Korea is not in the equipment. Rather, the difference lies in the use of 3D printers in the classroom, which is about three times higher there. Although there are currently positive correlations between 3D Printing Experts, classroom use and the availability of 3D printers, for some (particularly in the 3D Printing Novice group) the 3D printer remains unused despite its availability. In a study by Drossel et al. (2017), the availability of sufficient ICT equipment was a significant factor for the integration of computers in only one of three countries. Due to the different ways in which 3D Printing can be integrated into the classroom (even outsourcing printing is possible; Kantaros et al., 2022), the availability of equipment is not necessarily an essential factor. In our study, however, there is a strong correlation between the availability of 3D printers and their integration into the classroom. Furthermore, access to 3D printers is not exceptionally low compared to data describing the accessibility of tablet sets to whole classes, which is reported at 66% for Germany (IU Internationalen Hochschule, 2022). The fact that only 22% of teachers have already integrated 3D Printing into the classroom suggests that the general availability of printers may not be the limiting factor, as 45% of teachers report having one in their school. So there have to be other factors, e.g., interest or motivational aspects.

The interest in 3D Printing is quite positive in the sample of teachers with different subjects and 3D Printing knowledge studied here, which corresponds to the “desire” to acquire 3D printers at their own school. However, the fact that around half of teachers report a high or very high level of interest in attending a training course may indicate that they feel insecure in some way. Access to technology can be one of the many barriers teachers face when planning and implementing digitally-enhanced lessons (e.g., Pelgrum, 2001). However, even if access were a prerequisite for engaging with technology, many other factors or barriers come into play that do not usually resolve themselves (Hew and Brush, 2007). Thus, the availability of 3D printers in schools does not (consistently) lead to their integration into the classroom. As the data shows, the use of 3D Printing in the private sector is already more pronounced in all STEM groups. Therefore, experience gained in this area may support integration into the classroom in the future. Both the range of instructional materials/concepts and, in particular, the range of training to build competencies and self-efficacy are at least equally important as the equipment for integrating 3D Printing into one’s teaching (e.g., Arslan and Erdogan, 2021). The adoption of novel technology is largely determined by personal factors. Performance expectancy (related to advantages of 3D Printing), anxiety (of making mistakes or against 3D Printing technology), and attitudes toward technology use are significant predictors of teachers’ behavioral intentions when using new technologies (Holzmann et al., 2020), as 3D Printing represents for many. Training enables teachers to first gain their own experience with the technology as learners, to reflect on its pedagogical value, to form positive attitudes and reduce fears, and then to learn as teachers how to use 3D models in the classroom (e.g., Novak et al., 2021; Chen et al., 2023). Among the teachers in our study, the completion of a training course on 3D Printing is clearly underrepresented. The reasons given for this are a lack of interest and time, as well as a lack of courses on offer. While the intrinsic motivation to participate must be provided by the teachers themselves, the findings point to necessary implications in pre-service and in-service teacher training. Especially in the regular school routine of a working teacher, a lot can be achieved with short one-day training courses on 3D Printing, especially when time is a barrier (Novak, 2019). But that is also the case for courses at universities as shown by Ishutov et al. (2021) or Thoms et al. (2022). In interpreting our data in this context, it is worth noting that the proportion of newcomers is increasing from technology through computing to 3D Printing knowledge. 3D Printing as a technology in the area of modeling and simulation is by far the area with the highest demand for or low supply of training (in Germany: Diepolder et al., 2021).

5.3. Perceptions of competence development with 3D Printing in the classroom (Q2)

The teachers in this study perceived an increase in competencies through the integration of 3D Printing, especially in the areas of creativity, modeling and technology, closely followed by problem solving and scientific inquiry. This goes hand in hand with teacher/educator and student competencies (Trust and Maloy, 2017; Assante et al., 2020), but can be further differentiated in terms of 3D Printing expertise for the present study. We found empirical support for one of our three research hypotheses. The presumed differences depending on the expert status of the teachers with regard to 3D Printing can be partially confirmed in the areas of competence development through the incorporation of 3D Printing investigated here (H2.3). Teachers with a high level of 3D Printing knowledge rate 3D Printing as a valuable tool for developing competencies in scientific inquiry, problem solving and general aspects of digitalization. For fostering competencies in scientific inquiry, problem solving there is also a high correlation with 3D Printing knowledge. Since a comparison of teachers with high and low technological and computer science knowledge does not show significant results, it seems that knowledge of 3D Printing in particular is required to gain this insight, at least in theory. Even more interesting is the fact that the experience of integrating 3D Printing into the classroom does not seem to change these assumptions related to such areas of competence development. In contrast, teachers who have already integrated 3D Printing into their teaching rate the same items significantly higher, but seem to see further potential in additional areas such as social and communication competencies. It seems that seeing students working in the field of 3D Printing enables teachers to identify potential that cannot be derived from theoretical reflection alone. This is in line with the findings of Thyssen et al. (2021), who show that willingness and plans to use ICT in the future show stronger correlations with their current use than with Technological Knowledge or Technological Pedagogical Knowledge (according to the TPACK model, Koehler et al., 2013), and even has a higher weight as a predictor in a regression model.

With regard to the three age groups and three subgroups on the number of STEM subjects formed in this study, a presumed relationship with the assessment of the development of competences with 3D Printing cannot be established (H2.1 and H2.2 must be rejected for the present sample).

5.4. Perceptions about the integration of 3D Printing in the classroom (Q3)

In addition to the external barriers, which the teacher has little control over, they themselves still face the challenge of thinking about the benefits of digital technologies and deriving potentials (possibly also 3D Printing in combination with other digital technologies, Caldarone, 2020) as well as acquiring knowledge and skills for integrating them [second-order barriers according to Ertmer (1999)]. Our data show that knowledge of 3D Printing and a differentiated, rather than general, approach seem to have an impact on the assessment of didactic and methodological aspects that are essential for considering the benefits of using 3D Printing in the classroom. Both factors led to significant differences when comparing the 3D Printing expertise groups and the way of rating, respectively. In particular, the increase in the rating of methodological integration when assessed in a more nuanced approach using a multiple item scale could indicate that practical trainings that allow experiencing real lessons and thus actual methodological implementations could have beneficial effects due to the observations that can be made. This type of peer observation is helpful in several variations (Hamilton, 2013). Furthermore, more than 50% of the items tested for methodological and/or didactic integration of 3D Printing correlate significantly with knowledge of 3D Printing. Taken together, this demonstrates the need for more training courses in which teachers can acquire the relevant knowledge and adequate perspectives on 3D Printing. It would be important to provide teachers (including pre-service teachers) with approaches for differentiated consideration, pedagogical concepts and models for integrating 3D Printing in the classroom; such as essential content of in-service training (Assante et al., 2020).

Meier et al. (2022) provide a theoretical approach that reflects priorities with respect to didactic or methodological, product or process oriented, and subject-internal aspects. Based on these three perspectives of integration in the classroom, a three-dimensional space can be created to reflect on the objectives and content of the lesson. The base is formed by two fixed axes that can represent the focus of integration in terms of didactics or methodology (axis 1 in Figure 1), and the printer itself, the printed product, or the process in between (axis 2 in Figure 1). The 3D printer as a physical device and the 3D Printing product is very present in the perceptions of the surveyed teachers in the present study. In contrast, the inclusion or the perception of the design process as a possible way to integrate 3D Printing in the classroom is more in the middle range of agreement. This could be due to the fact that most of the teaching concepts and materials available focus often on a specific 3D Printing product (e.g., Jones and Spencer, 2018; Haverkamp et al., 2021). In addition, perceptions of the 3D printed product and its use in the classroom are probably closer to the common use of media (in this case, models) in the subject lesson. In contrast, the integration of the design process for printing is linked to knowledge of the technology and the process steps and usually also leads to changes in the teaching concept.

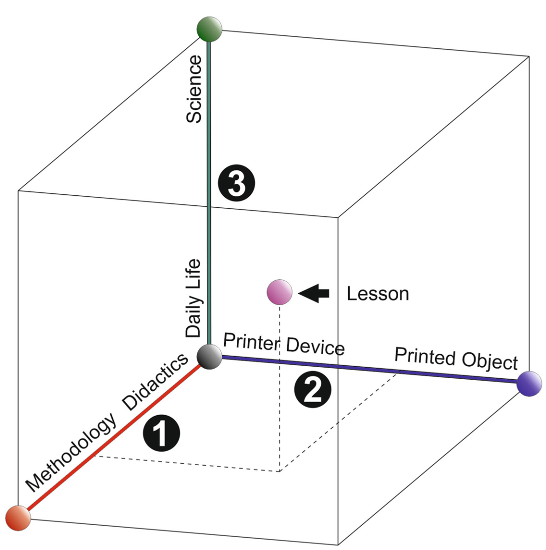

FIGURE 1
Dimensions and decision fields of the integration of 3D Printing during lesson planning, adapted from Meier et al. (2022).

Learning situations in which students digitally design models themselves and then physically print them out are not possible without partially adopting concepts from the maker movement, and are closely linked to a constructivist understanding of learning (Pearson and Dubé, 2022). In the creation of self-directed learning environments in which individual and differentiated approaches to learning are made possible, there is potential for the 3D Printing Experts in this study in learning through 3D Printing. They differed significantly from the 3D Printing Novices in their conception of this. This observation could be interpreted to mean that 3D Printing supports constructivist learning or approaches that incorporate design thinking concepts or methods based on them and elements derived from them are seen as promising by teachers. The extent to which these teachers also methodically implement design thinking supported by 3D Printing into their own classrooms can vary widely and does not necessarily need to take advantage of the full potential of 3D Printing. In fact, it may be as simple as just integrating a few elements (Leinonen et al., 2020). However, this was not explicitly addressed in the context of the study or covered with specific items and should be explored in more detail in follow-up studies.

Perpendicular to axes 1 and 2 is a third, context-dependent axis (axis 3 in Figure 1), each consisting of a pair of terms describing relevant contextual areas. The vertical axis is to be understood as a flexible set of, possibly subject-dependent axes to capture the relevant contexts. In STEM education, contexts can be represented by axes with different extremes, such as science or everyday life or for other subjects and contexts 3D Printing/design process and 3D Printing equipment technology, chemistry of 3D Printing and technology of 3D Printing process. The comparison of our data, according to which STEM teachers see a higher possibility of integrating 3D Printing as content in the sense of a corresponding science technology in their subjects, with the presented model (Figure 1, axis 3: daily life/science) allows two interesting interpretations: the model can (a) explain the differences between STEM subjects that have emerged on this topic and (b) potentially predict a larger space for the integration of 3D Printing in STEM subjects. For contextual perspective and adaptation, different pairs of terms should be formulated for the third axis depending on the subject. STEM subjects or teachers' perceptions do not seem to differ fundamentally in areas relevant to lesson planning in general. The perception of the possibilities of methodological or didactic integration and the integration of printed products or the design process do not seem to be STEM specific. This means that a model with a more or less general but adaptable structure may be appropriate and flexible enough to account for the observed differences. The needs of different subjects can be met by adjusting the third axis for analytical purposes. However, this will not change the observation that at least right now STEM subjects have the potential of integrating 3D Printing in the context of science (e.g., HU and Jiang, 2017; Walker and Humphries, 2019) matching higher assessment of an integration in SD concepts while teachers of other subjects seem to assess reduced possibilities for both fields.

5.5. Link to the (NON-)STEM subjects taught

For self-reported knowledge in the mentioned areas (Table 2), a significant difference can be found between teachers with two STEM subjects and teachers with one or no STEM subject. One explanation for these differences may be the specificity of computer science and 3D Printing knowledge in particular, which may be associated with related technologies in science. This interpretation would be supported by the finding that teachers with two STEM subjects, when they think of 3D Printing, are less likely to think of it as a field for other subjects, and see opportunities for integration into the classroom as an established technology in scientific fields corresponding to their subjects. The use of digital technologies is less influenced by the subject in terms of scope, but is certainly influenced by the subject in terms of the design and type of technologies incorporated (e.g., Záhorec et al., 2019). However, the number of putative effects that may exist in terms of the number of STEM subjects taught is small. Apart from the actual use of 3D printers, significant effects can only be found for 6 items, three of which, as reported, concern the information on the existing knowledge, two the reference to the own teaching subjects and one the promotion of sustainability competences with corresponding SD concepts using 3D Printing. This suggests that there may also be determinants in the latter area, which are not directly linked to 3D Printing knowledge but to the number of STEM subjects compared to the differences found for stated knowledge. Similarly, the alignment between STEM and non-STEM teachers in the use of 3D Printing evidenced in other studies and countries (Chen et al., 2023) may also be evident in our study. Certainly, teachers' perceptions and perspectives will change, driven by self-taught dynamics or those specifically initiated by in-service training.

6. Implication: what can be derived from this study for future training concepts?

Teaching with 3D Printing in the classroom is now coming up against not so much equipment limitations as training limitations (Pearson and Dubé, 2022), which are narrowly defined by a (still) very small number. Following on from the reported findings on perceptions of 3D Printing in the present study, training courses for Novices and Experts need to be developed, adapted to the level of experience and knowledge as well as to the interests of the participants. Ideally, these courses should include a pedagogical approach to the use of 3D Printing in the classroom (Assante et al., 2020) and practical approaches in schools, rather than focusing solely on technical aspects. As it is clear that the use of 3D Printing in the classroom provides additional insights and perceptions in terms of fostering interaction and communication skills, new training approaches could also be considered. Implementations that allow teachers to observe real lessons and experience student interaction and communication could potentially provide such perceptions directly. New or more hands-on training formats raise questions about the impact and sustainability of training in technology use.

Another targeted alternative would be further training with observation of the teaching of experts, with novices even assisting as co-teachers after their own training. Such an approach would specifically encourage peer support, which is difficult to build due to the still small number of 3D Printing users. This concept could be used to initiate a specific form of cooperation between teachers, the Professional Learning Communities (PLC). In PLCs, ideally, practitioners ("teachers as learners"; Bonsen and Rolff, 2006, p. 169) work together continuously, cooperatively and critically by exchanging ideas about their own teaching and subject content. It is assumed that teachers' collaboration can support their professional development (e.g., Terhart and Klieme, 2006; Methlagl, 2022). The assumed positive relationship between teacher collaboration and teacher competence is derived from situated learning approaches (Putnam and Borko, 2000; Borko, 2004). Learning to teach in applied situations/contexts supports the transfer of "new learning" into one's own or future teaching.

7. Limitations and further research

Limiting factors for the validity of the findings in this study include the sample generation. With regard to the equipment with 3D printers and their use, the sample may be biased as several teachers from a school may have responded to the survey and it is unclear how many teachers were actually in the same school. However, the specification of the school could have had an unfavorable effect on the feeling of anonymity during participation and consequently lead to lower participation. As a result, the respondents were not asked to name their own school.

With the intention of broadly capturing and describing the teachers' perceptions, a questionnaire with content-rich items and different item formats was developed. This has a limiting effect on the evaluation procedures and the nature of the results. Models that explain the relationships and interactions of factors in an explanatory way cannot be derived from the data collected for a descriptive survey, as there are no scales for variables that (could) interact in model contexts. This is where future research is needed to develop appropriate scales (e.g., Gürer et al., 2019) to fit new or existing models on the basis of available data. In addition, qualitative methods should be increasingly included in the collection and analysis of attitudes and perceptions about 3D Printing. This is already more common in intervention or evaluation studies (e.g., Song, 2018), but could be expanded with an eye toward teachers' general and subject-specific conceptions of 3D Printing. This will also require comprehensive statistical surveys of 3D Printing, equipment, and existing training to validate our findings. In this context, an analysis of existing training concepts would be particularly helpful for the development of new training courses (e.g., Novak and Wisdom, 2020; Cuun, 2021).

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

Author contributions

CT and MM analyzed, revised, and discussed the data. All authors contributed to the article and approved the submitted version.

Acknowledgments

The successively developed questionnaire also builds on preliminary work by a group including the authors, Lars-Jochen Thoms, Thomas Schubatzky, and Markus Obczovsky. The data collection was done with the extensive help and organization of Agnetha Nißler.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ali, N., Khine, M. S., Wisdom, S., and Novak, E. (2019). "Using 3D printing to enhance STEM teaching and learning" in *Integrating 3D printing into teaching and learning. Practitioners' perspectives*. eds. N. Ali and M. S. Khine (Leiden: Brill), 187–205.

Arslan, A., and Erdogan, I. (2021). Use of 3D printers for teacher training and sample activities. *Int. J. Progress. Educ.* 17, 343–360. doi: 10.29329/ijpe.2021.346.22

Aslan, A., and Celik, Y. (2020). A literature review on 3D printing technologies in education. *Int J 3D Print Technol Digit Industry* 6, 592–613. doi: 10.46519/ij3dptdi.1137028

Assante, D., Cennamo, G. M., and Placidi, L. (2020). 3D printing in education: an European perspective. In: *IEEE Global Engineering Education Conference (EDUCON)* 1133–1138.

Augusto, I., Monteiro, D., Girard-Dias, W., dos Santos, T. O., Rosa Belmonte, S. L., Pinto de Oliveira, J., et al. (2016). Virtual reconstruction and three-dimensional printing of blood cells as a tool in cell biology education. *PLoS One* 11:e0161184. doi: 10.1371/journal.pone.0161184

Bagley, J. R., and Galpin, A. J. (2015). Three-dimensional printing of human skeletal muscle cells: An interdisciplinary approach for studying biological systems. *Biochem. Mol. Biol. Educ.* 43, 403–407. doi: 10.1002/bmb.20891

Barreto, J. E. F., Kubrusly, B. S., Lemos Filho, C. N. R., Silva, R. S., Façanha, S. d. O., Santos, J. C. C. d., et al. (2022). 3D printing as a tool in anatomy teaching. *Int J Innov Educ Res* 10, 58–71. doi: 10.31686/ijier.vol10.iss6.3771

Bécar, J. P., Vareille, J., Cayez, V., and Notteau, Y. (2017). 3D printing ist boosting the Student's creativity. In: *ICERI2017 Proceedings*, 1817–1823.

Berg, S., Ridenour Benz, C., Lasley, T. J., and Raisch, C. D. (1998). Exemplary technology use in elementary classrooms. *J. Res. Comput. Educ.* 31, 111–122. doi: 10.1080/08886504.1998.10782245

Bicer, A., Nite, S. B., Capraro, R. M., Barroso, L. R., Capraro, M. M., and Lee, Y. (2017). Moving from STEM to STEAM: the effects of informal STEM learning on students' creativity and problem solving skills with 3D printing. In: *IEEE Frontiers in Education Conference (FIE)* 1-6.

Bingimlas, K. A. (2009). Barriers to the successful integration of ICT in teaching and learning environments: a review of the literature. *Eurasia J Maths Sci Technol Educ* 5, 235–245. doi: 10.12973/ejmste/75275

Bonorden, M., and Papenbrock, J. (2022). Evidence-based optimization of classroom teaching units using 3D printers for designing models—from the 2D picture to the 3D flower model. *Educ. Sci.* 12:11. doi: 10.3390/educsci12110831

Bonsen, M., and Rolff, H.-G. (2006). Professionelle Lerngemeinschaften von Lehrerinnen und Lehrern (Professional learning communities for teachers). *Zeitschrift für Pädagogik* 52, 167–184. doi: 10.25656/01:4451

Borko, H. (2004). Professional development and teacher learning: mapping the terrain. *Educ. Res.* 33, 3–15. doi: 10.3102/0013189X033008003

Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion* (3rd Edn.). München: Pearson.

Caldarone, A. (2020). Mixed reality, 3D printing, and storytelling: methodologies for the creation of multi-sensory scenarios in the field of cultural heritage. In: *Proceedings CAA*.

Chen, Y., Cao, L., and Zhang, Y. (2023). Teachers as makers: how K-12 teachers design 3D making lessons for classroom teaching. *Educ. Inf. Technol.* 28, 6947–6975. doi: 10.1007/s10639-022-11475-w

Chen, J., and Cheng, L. (2021). The influence of 3D printing on the education of primary and secondary school students. *J. Phys. Conf. Ser.* 1976:012072. doi: 10.1088/1742-6596/1976/1/012072

Chien, Y.-H., and Chu, P.-Y. (2018). The different learning outcomes of high school and college students on a 3D-printing STEAM engineering design curriculum. *Int. J. Sci. Math. Educ.* 16, 1047–1064. doi: 10.1007/s10763-017-9832-4

Choi, H., and Kim, J. M. (2018). Implications for activating 3D printer use for education in elementary and secondary schools. *Int J Adv Sci Engineer Inform Technol* 8, 1546–1551. doi: 10.18517/ijaseit.8.4-2.5722

Cohen, J. (1992). A power primer. *Psychol. Bull.* 112, 155–159. doi: 10.1037/0033-2909.112.1.155

Cuun, H. (2021). A study on the impact of 3D printing and artificial intelligence on education and learning process. *Sci. Prog.* 2021, 1–5. doi: 10.1155/2021/2247346

Diepolder, C., Weitzel, H., Huwer, J., and Lukas, S. (2021). Verfügbarkeit und Zielsetzungen digitalisierungsbezogener Lehrkräftefortbildungen für naturwissenschaftliche Lehrkräfte in Deutschland. *ZfDN* 27, 203–214. doi: 10.1007/s40573-021-00134-1

Drossel, K., Eickelmann, B., and Gerick, J. (2017). Predictors of teachers' use of ICT in school – the relevance of school characteristics, teachers' attitudes and teacher collaboration. *Educ. Inf. Technol.* 22, 551–573. doi: 10.1007/s10639-016-9476-y

Ertmer, P. A. (1999). Addressing first- and second-order barriers to change: strategies for technology integration. *Educ. Technol. Res. Dev.* 47, 47–61. doi: 10.1007/BF02299597

Ertmer, P. A., Ottenbreit-Leftwich, A. T., and Tondeur, J. (2015). Teachers' beliefs and uses of technology to support 21st-century teaching and learning. *Int. Handbook Res. Teach Beliefs* 403, 403–419.

Ford, S., and Minshall, T. (2019). Invited review article: where and how 3D printing is used in teaching and education. *Addit. Manuf.* 25, 131–150. doi: 10.1016/j.addma.2018.10.028

Greenhalgh, S. (2016). The effects of 3D printing in design thinking and design education. *J Engineer Design Technol* 14, 752–769. doi: 10.1108/JEDT-02-2014-0005

Guo, R. X., Dobson, T., and Petrina, S. (2008). Digital natives, digital immigrants: An analysis of age and Ict competency in teacher education. *J. Educ. Comput. Res.* 38, 235–254. doi: 10.2190/EC.38.3.a

Gürer, M. D., Tekinarslan, E., Kocaayak, I., and Gonultas, S. (2019). Development and validation of an attitude assessment scale for the use of 3D printing in education. *Int. J. Educ. Dev. Using Inf. Commun. Technol.* 15:204.

Hamilton, E. R. (2013). His ideas are in my head: peer-to-peer teacher observations as professional development. *Prof. Dev. Educ.* 39, 42–64. doi: 10.1080/19415257.2012.726202

Haverkamp, N., Havemann, J., Holz, C., Ubben, M., Schlummer, P., and Pusch, A. (2021). A new implementation of Kundt's tube: 3D-printed low-cost set-up using ultrasonic speakers. *Phys. Educ.* 56:025023. doi: 10.1088/1361-6552/abd0d7

Hew, K. F., and Brush, T. (2007). Integrating technology into K-12 teaching and learning: current knowledge gaps and recommendations for future research. *Educ. Technol. Res. Dev.* 55, 223–252. doi: 10.1007/s11423-006-9022-5

Holzmann, P., Schwarz, E. J., and Audretsch, D. B. (2020). Understanding the determinants of novel technology adoption among teachers: the case of 3D printing. *J. Technol. Transf.* 45, 259–275. doi: 10.1007/s10961-018-9693-1

Hsu, Y.-C., Baldwin, S., and Ching, Y.-H. (2017). Learning through making and maker education. *Tech Trends* 61, 589–594. doi: 10.1007/s11528-017-0172-6

Hu, L., and Jiang, G. (2017). 3D printing techniques in environmental science and engineering will bring new innovation. *Environ. Sci. Technol.* 51, 3597–3599. doi: 10.1021/acs.est.7b00302

Ishutov, S., Hodder, K., Chalaturnyk, R., and Zambrano-Narvaez, G. (2021). A 3D printing short course: a case study for applications in the geoscience teaching and communication for specialists and non-experts. *Front. Earth Sci.* 9:601530. doi: 10.3389/feart.2021.601530

Isi, C., and Gurley, M. (2023). 3D PrintingTrend report 2023. Market insights and forecasts in additive manufacturing. HUBS. Available at: https://www.Hubs.Com/get/trends/ (accessed 24 May 2023).

IU Internationalen Hochschule (2022). Wie digital sind unsere Schulen? Available at: https://static.iu.de/studies/Digitale_Bildung_Studie.pdf (Accessed 31 May 2023).

Jones, O. A. H., and Spencer, M. J. S. (2018). A simplified method for the 3D printing of molecular models for chemical education. *J. Chem. Educ.* 95, 88–96. doi: 10.1021/acs.jchemed.7b00533

Kalaskar, D. M. (2022). *3D printing in medicine*. *2nd* Edn Sawston: Elsevier Science, Woodhead Publishing.

Kantaros, A., Diegel, O., Piromalis, D., Tsaramirsis, G., Khadidos, A. O., Khadidos, A. O., et al. (2022). 3D printing: making an innovative technology widely accessible through makerspaces and outsourced services. *Mater. Today* 49, 2712–2723. doi: 10.1016/j.matpr.2021.09.074

Koehler, M. J., Mishra, P., and Cain, W. (2013). What is technological pedagogical content knowledge (TPACK)? *J. Educ.* 193, 13–19. doi: 10.1177/002205741319300303

Kwon, H. (2017). Effects of 3D printing and design software on students' interests, motivation, mathematical and technical skills. *J STEM Educ* 18, 37–42.

Kwon, S.-H., Lee, Y.-J., and Kwon, Y.-J. (2020). An active learning approach to investigate the ecosystem of tide flats using 3D modeling and printing. *J. Biol. Educ.* 54, 88–97. doi: 10.1080/00219266.2018.1546760

Leinonen, T., Virnes, M., Hietala, I., and Brinck, J. (2020). 3D printing in the wild: adopting digital fabrication in elementary school education. *Int J Art Design Educ* 39, 600–615. doi: 10.1111/jade.12310

Lorenz, R., and Eickelmann, B. (2022). "Nutzung digitaler Medien im Unterricht der Sekundarstufe I und Nutzungsbedingungen im Trendvergleich von 2017 und 2021" in *Schule digital – der Länderindikator 2021. Lehren und Lernen mit digitalen Medien in der Sekundarstufe I in Deutschland im Bundesländervergleich und im Trend seit 2017*. eds. R. Lorenz, S. Yotyodying, B. Eickelmann and M. Endberg (Münster, New York: Waxmann), 63–88.

Martin, L. (2015). The promise of the maker movement for education. *J Pre-Coll Engineer Educ Res* 5, 30–39. doi: 10.7771/2157-9288.1099

Martinez, S. L., and Stager, G. (2013). *Invent to learn: Making, tinkering, and engineering in the classroom*. Torrance: Constructing Modern Knowledge Press.

Medina Herrera, L., Castro Pérez, J., and Juárez Ordóñez, S. (2019). Developing spatial mathematical skills through 3D tools: augmented reality, virtual environments and 3D printing. *Int. J. Interact. Des. Manuf.* 13, 1385–1399. doi: 10.1007/s12008-019-00595-2

Meier, M., Schubatzky, T., Obczovsky, M., Thoms, L.-J., and Thyssen, C. (2022). Fachdidaktische Perspektiven und Szenarien des 3D-Drucks im naturwissenschaftlichen Unterricht. *MNU J.* 1, 79–84.

Meier, M., and Thyssen, C. (2021). Vom Original zum 3D-Objekt. Dimensionen und Guidelines zum 3D-Druck im Unterricht. *Unterricht Biologie, Reihe BioDigital.* 463, 43–47.

Meskill, C., Mossop, J., and DiAngelo, S., nd Pasquale, R. K. (2002). Expert and novice teachers talking technology: precepts, concepts, and misconcepts. *Lang. Learn. Technol.* 6, 46–57.

Methlagl, M. (2022). Patterns of teacher collaboration, professional development and teaching practices: a multiple correspondence analysis of TALIS 2018. *Int J Educ Res Open* 3:100137. doi: 10.1016/j.ijedro.2022.100137

Monkovic, J. M., Jones, S. M., Nicolas, M., Katyal, P., Punia, K., Noland, D., et al. (2022). From concept to reality: the use and impact of 3D prints as academic tools for high school biology education. *J. Biol. Educ.* 56, 528–539. doi: 10.1080/00219266.2020.1858927

Munaz, A., Vadivelu, R. K., John, J. S., Barton, M., Kamble, H., and Nguyen, N.-T. (2016). Three-dimensional printing of biological matters. *J Sci Adv Materials Devices* 1, 1–17. doi: 10.1016/j.jsamd.2016.04.001

Novak, J. I. (2019). "Re-educating the educators: collaborative 3D printing education" in *Interdisciplinary and international perspectives on 3D printing in education*. eds. I. Santos, N. Ali and S. Areepattamannil (Hershey: IGI Global), 28–49.

Novak, E., Brannon, M., Librea-Carden, M. R., and Haas, A. L. (2021). A systematic review of empirical research on learning with 3D printing technology. *J. Comput. Assist. Learn.* 37, 1455–1478. doi: 10.1111/jcal.12585

Novak, E., and Wisdom, S. (2020). "Using 3D printing in science for elementary teachers" in *Active learning in college science: The case for evidence-based practice*. eds. J. J. Mintzes and E. M. Walter (Cham: Springer), 729–739.

Oss Boll, H., de Castro Leitão, M., Garay, A. V., Batista, A. C. C., de Resende, S. G., da Silva, L. F., et al. (2023). SynBio in 3D: the first synthetic genetic circuit as a 3D printed STEM educational resource. *Front. Educ.* 8:1110464. doi: 10.3389/feduc.2023.1110464

Pantazis, A., and Priavolou, C. (2017). 3D printing as a means of learning and communication: the 3Ducation project revisited. *Telematics Inform.* 34, 1465–1476. doi: 10.1016/j.tele.2017.06.010

Pearson, H. A., and Dubé, A. K. (2022). 3D printing as an educational technology: theoretical perspectives, learning outcomes, and recommendations for practice. *Educ. Inf. Technol.* 27, 3037–3064. doi: 10.1007/s10639-021-10733-7

Pelgrum, W. J. (2001). Obstacles to the integration of ICT in education: results from a worldwide educational assessment. *Comput. Educ.* 37, 163–178. doi: 10.1016/S0360-1315(01)00045-8

Pernaa, J., and Wiedmer, S. (2020). A systematic review of 3D printing in chemistry education – analysis of earlier research and educational use through technological pedagogical content knowledge framework. *Chem Teach Int* 2:20190005. doi: 10.1515/cti-2019-0005

Putnam, R. T., and Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educ. Res.* 29, 4–15. doi: 10.3102/0013189X029001004

Renner, M., and Griesbeck, A. (2020). Think and print: 3D printing of chemical experiments. *J. Chem. Educ.* 97, 3683–3689. doi: 10.1021/acs.jchemed.0c00416

Schelly, C., Anzalone, G., Wijnen, B., and Pearce, J. M. (2015). Open-source 3-D printing technologies for education: bringing additive manufacturing to the classroom. *J Vis Lang Comput* 28, 226–237. doi: 10.1016/j.jvlc.2015.01.004

Sfetsas, T., Patsatzis, S., and Chioti, A. (2021). A review of 3D printing techniques for bio-carrier fabrication. *J. Clean. Prod.* 318:128469. doi: 10.1016/j.jclepro.2021.128469

Song, M. J. (2018). Learning to teach 3D printing in schools: how do teachers in Korea prepare to integrate 3D printing technology into classrooms? *Educ. Media Int.* 55, 183–198. doi: 10.1080/09523987.2018.1512448

Sugar, W., Crawley, F., and Fine, B. (2004). Examining teachers' decisions to adopt new technology. *Educ. Technol. Soc.* 7, 201–213.

Terhart, E., and Klieme, E. (2006). Kooperation im Lehrerberuf: Forschungsproblem und Gestaltungsaufgabe. *Zeitschrift für Pädagogik* 52, 163–166. doi: 10.25656/01:4450

Thoms, L. J., Hoyer, C., and Girwidz, R. (2022). "A teacher training course on using digital Media for Acquisition, visualization and 3D printing of complex data and for fostering pupils' experimental skills" in *Physics teacher education. Challenges in physics education*. eds. J. Borg Marks, P. Galea, S. Gatt and D. Sands (Cham: Springer)

Thyssen, C., Pankow, A., Klaeger, K., and Chernyak, D. (2021). Kompetenzen und Nutzungsperspektiven im Bereich digitaler Medien bei Lehrkräften zur Erkenntnisgewinnung in den Naturwissenschaften. *Zeitschrift Empirische Pädagogik* 35, 112–135.

To, T. T., Al Mahmud, A., and Ranscombe, C. (2023). Teaching sustainability using 3D printing in engineering education: An observational study. *Sustainability* 15:7470. doi: 10.3390/su15097470

Trust, T., and Maloy, R. W. (2017). Why 30 print? The 21st-century skills students develop while engaging in 3D printing projects. *Comput. Sch.* 34, 253–266. doi: 10.1080/07380569.2017.1384684

Walker, M., and Humphries, S. (2019). 3D printing: applications in evolution and ecology. *Ecol. Evol.* 9, 4289–4301. doi: 10.1002/ece3.5050

Wang, L., Luo, J., An, L., Zhou, X., Yin, C., and Ma, H. (2021). Promoting junior school Students' spatial ability through 3D printing. In: *2021 Tenth International Conference of Educational Innovation through Technology (EITT)* 100–105.

Wilk, R., Likus, W., Hudecki, A., Syguła, M., Różycka-Nechoritis, A., and Nechoritis, K. (2020). What would you like to print? Students' opinions on the use of 3D printing technology in medicine. *PLoS One* 15:e0230851. doi: 10.1371/journal.pone.0230851

Yan, Q., Dong, H., Su, J., Han, J., Song, B., Wei, Q., et al. (2018). A review of 3D printing technology for medical applications. *Engineering* 4, 729–742. doi: 10.1016/j.eng.2018.07.021

Ye, Z., Dun, A., Jiang, H., Nie, C., Zhao, S., Wang, T., et al. (2020). The role of 3D printed models in the teaching of human anatomy: a systematic review and meta-analysis. *BMC Med. Educ.* 20:335. doi: 10.1186/s12909-020-02242-x

Ye, Z., Jiang, H., Bai, S., Wang, T., Yang, D., Hou, H., et al. (2023). Meta-analyzing the efficacy of 3D printed models in anatomy education. *Front. Bioeng. Biotechnol.* 11:1117555. doi: 10.3389/fbioe.2023.1117555

Yıldırım, G. (2018). Teachers' opinions on instructional use of 3D printers: a case study. *Int Online J Educ Sci* 10, 304–320. doi: 10.2139/ssrn.3331158

Záhorec, J., Nagyová, A., and Hašková, A. (2019). Teachers' attitudes to incorporation digital means in teaching process in relation to the subjects they teach. *Int J Engineer Pedag* 9, 100–120. doi: 10.3991/ijep.v9i4.11064

Zhou, L.-Y., Fu, J., and He, Y. (2020). A review of 3D printing technologies for soft polymer materials. *Adv. Funct. Mater.* 30:2000187. doi: 10.1002/adfm.202000187

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 07 September 2023
DOI 10.3389/feduc.2023.1232423

OPEN ACCESS

EDITED BY
Wang-Kin Chiu,
The Hong Kong Polytechnic University, China

REVIEWED BY
Maria Beatrice Ligorio,
University of Bari Aldo Moro, Italy
Milan Kubiatko,
J. E. Purkyne University, Czechia

*CORRESPONDENCE
Niina Halonen
✉ niina.halonen@helsinki.fi

RECEIVED 31 May 2023
ACCEPTED 08 August 2023
PUBLISHED 07 September 2023

CITATION
Halonen N, Ståhle P, Juuti K, Paavola S and Lonka K (2023) Catalyst for co-construction: the role of AI-directed speech recognition technology in the self-organization of knowledge. *Front. Educ.* 8:1232423. doi: 10.3389/feduc.2023.1232423

Catalyst for co-construction: the role of AI-directed speech recognition technology in the self-organization of knowledge

Niina Halonen [1]*, Pirjo Ståhle[2], Kalle Juuti[1], Sami Paavola[1] and Kirsti Lonka[1,3]

[1]Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland, [2]School of Engineering, Aalto University, Espoo, Finland, [3]Optentia Research Focus Area, North-West University, Vanderbijlpark, South Africa

The purpose of this study was to analyze knowledge co-construction as a self-organization process and the role of technology as its catalyst. Novel AI-directed speech recognition technology and the artifacts it generates were deployed to scaffold the knowledge co-construction process in two groups of pre-service teachers in a science education context. Throughout the lesson, the focus of the learning tasks was on pedagogical content knowledge and students' preconceptions. Analysis was conducted through the key characteristics of the social system's self-organization theory. The process of self-organization refers to the system's capacity to diverge from familiar structures, perspectives, and operations. Through the lenses of system theories, the active role of artifacts in co-construction was grasped and the role of technology in the self-organization of knowledge was analyzed. The pedagogical design of knowledge co-construction followed the principles of student-engaging learning. The technology used in co-construction was novel speech recognition AI software, which produced visual and editable word cloud artifacts from oral discussions on the large-format screen to edit. The data included videos and audio recordings. In this qualitative study, a content analysis and interaction analysis were used with descriptive analysis. The results showed that when technology became visible, as an active component of the system, artifacts triggered key signs of the social system's self-organization in co-construction. Exchange of information, "entropy levels," were rapidly increased, and different viewpoints were expressed. Also, "chaos zones," far-from-equilibrium states, were reached in both groups. Editable artifacts on the screen represented bifurcation spaces where groups' discussions were crystallized for the first time. Information was further categorized and evaluated through artifacts and this demonstrated how the groups processed communication into learning insights. Based on the results, the role played by this kind of technology was significant in the self-organization of knowledge. Materialized artifacts pushed the groups from small group conversation phases, comfort zones, toward uncertainty and confusion, which are central in self-organization. Technology in the system is seen not only as an interactor but also as an active agent that can facilitate epistemic emotions and support the group in the self-organization of knowledge.

KEYWORDS

knowledge co-construction, self-organization of knowledge, systems theoretic approach, technology-enhanced learning, artificial intelligence, speech recognition, technology-mediated learning

1. Introduction

In this study, the social system's self-organization theory approach (Ståhle, 1998) is adopted to analyze knowledge co-construction processes where technology is seen as one central component of the system. Through the lenses of system theories, it is possible to grasp the active role of artifacts in co-construction and analyze the role played by AI-directed speech recognition technology in the self-organization of knowledge. The system theory approach has not been widely used for analyzing co-construction processes. Group agency is widely studied (Dillenbourg et al., 2009; List and Pettit, 2011; Stenalt, 2021; Brod et al., 2023) but there are only a few empirical studies on group self-organizing processes of knowledge in the context of educational technologies (Scheel et al., 2022). Sawyer (2006, 2009) has connected a systemic approach to collaboration and formed the concept of collaborative emergence. Sawyer introduces the phenomena, studied by social scientists, that emerge from "complex systems of individuals in interaction." Ritella and Hakkarainen (2012) see technology-mediated learning practices as a similar phenomenon, a distributed system involving inter-psychological (social) and intra-psychological (individual) levels, materially embodied artifacts, and different perceptions of time.

Productive collaborative processes share similarities with social systems' self-organization processes of knowledge, which can be also equated with innovative learning and the emergence of new knowledge (Ståhle et al., 2020). If the indications of self-organization exist when a group of learners constructs knowledge, the process includes affordances to support higher-level learning and deeper cognitive processes (Bloom et al., 1956), which are also the aims of novel pedagogical approaches and 21st-century competencies. The challenge in existing pedagogical practices is still to generate higher-order cognitive processes, even though novel curriculums' content in recent years has shifted learning objectives from the idea of information transfer to higher-level learning (Härkki et al., 2021) such as the skills to analyze and solve problems and apply complex ideas (Haataja et al., 2023). The demands and aims of 21st-century education also challenge the pedagogical practices of using digital tools (Schleicher, 2018; Sanina et al., 2020).

The importance of collaboratively developed artifacts is highlighted in idea development and productive collaboration (Hennessy and Murphy, 1999; Barron, 2003; Paavola and Hakkarainen, 2009; Kangas et al., 2013). Also, in more recent learning research, Sawyer (2022) highlights the role of material tools in learning and collaborative creativity. He points out that artifact agency is a "positive force that drives the creative process forward." Jointly constructed artifacts have an agency that supports creativity.

In this study, a system refers to an entity composed of a student group, technology, artifacts produced by students and technology, and co-construction activities. In systems where knowledge is created by self-organization, the system must undergo iterative processes that involve feedback loops. Through iterative interactions between its components, a system is driven to a state of instability through which unexpected paths are opened for its development. These paths can format new structures, new knowledge, and higher-level learning outcomes. In the context of social systems, "entropy" is considered a key characteristic in generating iteration. The concept of entropy, used in this study (based on Ståhle, 1998) refers to the rich exchange of information, as well as increased disorder and randomness in a system. Self-organized systems must have the ability to produce and increase entropy: to discuss the topic from different points of view and elaborate on the information exchanged. An increase in entropy leads the system from its balance, "equilibrium," toward a far-from-equilibrium state in which confusion and uncertainty are necessary elements for knowledge to be analyzed, evaluated, and finally crystallized. When a system can operate with increased entropy in a far-from-equilibrium state, classifying and reflecting on information and tolerating challenges, the system has the opportunity to evolve and innovate a new order.

The self-organizing process of knowledge differs, for example, from a linearly scripted learning process, where the learning path is defined as precisely as possible. In a system's behavior, there must be room for uncertainty which creates space and opportunities for innovation through self-organization of knowledge. A self-organizing system has the freedom to act on and influence its decisions. Small changes can lead to significant shifts in the system's patterns and structures through *bifurcation moments*. The new order emerges suddenly from the system itself. Bifurcation moments in a broader sense can lead to significant shifts in learners' understanding and knowledge structures.

This study aims to analyze the knowledge co-construction process and the role of technology in the self-organization of knowledge. The research interest lies especially in technology as one active component of the system in co-construction, besides the group of learners. By taking a social system theory approach to the analysis of knowledge co-construction processes, scaffolded by AI-directed speech recognition technology, patterns of the key components that enable a group to self-organize their learning together are delineated.

From this context, the following research question arises:

> *"What is the role of AI-directed speech recognition technology in the co-construction process in terms of self-organization of knowledge?"*

In the following sections, the theoretical background is initially clarified. The main concepts of social self-organizing systems are revisited and our conceptual and analytical approach, reflecting learning theories, is framed. The research setting is then introduced, covering the technology-mediated co-construction of preservice teachers in science education and an empirical analysis of the case study. Finally, in the discussion, the focus is on evaluating the theoretical insights with empirical data aiming to identify the key components that enable a group to self-organize in knowledge co-construction.

2. Theoretical background

2.1. What is a system?

In systems thinking, the concept of a *system* refers to "a holistic entity composed of and dependent on a series of interconnected and interacting parts" (Ståhle et al., 2020, p. 191). Salazar (2002) defines a system in social sciences as a collection of mutually dependent elements that form a cohesive entity. A group can be construed as a system, given that it is comprised of mutually dependent elements, which may be conceptualized in various forms (such as group members, behaviors, and interactions) that contribute to a unified whole.

Systems have been studied from a variety of perspectives. Ståhle (1998, 2008) has distinguished three different research generations of systems paradigms, the first of which focused on systemic order and its predictability, the second on systems openness and steadiness, and the third on self-organizing systems. The systems thinking trend from the 1960s shifted research toward the unpredictable dynamics of systems, disorder, and the relationship between chaotic behavior and the emergence of order. These new viewpoints led to a new research approach, known as complexity theory (CT) or complex adaptive systems (CAS) theory (Ståhle, 2008). The field of complexity science refers to a growing body of research on dynamic non-linear feedback systems and self-organizing systems. Not only individuals but interconnected and interacting parts form a dynamic system whose ability to self-organize requires a chaotic state. In general, complexity sciences investigate the zone between order and chaos, where a system transitions to exhibiting a significant degree of dynamism, represented by an increased variety in the behavior of its constituent elements. The term "complexity" refers to the substantial interdependence among a system's elements, as well as a high degree of variety in their respective behaviors (Salazar, 2002).

2.2. Concept of self-organization

The process of self-organization refers to the ability of a system to move away from familiar structures, perspectives, and operations. The concept refers to the capacity of a system to create its own structure (mentally, socially, and physically) and to set rules that facilitate collaborative behavior without requiring external top-down control (Mitchell, 2009). Self-organization in a system may be said to occur when a system seemingly spontaneously develops new structural features and a new order after having progressed through a disruption. The disruption causes a kind of "crisis" in the system which moves it away from equilibrium, from its stable state. This movement is characterized by a display of a greater variety of behaviors than was the case when the system was functioning at, or close to, its equilibrium point (Salazar, 2002).

Prigogine's research can be said to be the most important contribution to the self-organization and dynamic systems paradigm. Prigogine, a famous chemist, pointed out in his theory of dissipative structures that physical or chemical systems appear to develop order out of chaos. Prigogine discovered new laws of nature that could connect the natural sciences to the human sciences, and he maintained that these laws are universal, and thus also applicable to social systems (Prigogine, 1976, p. 120–126; Ståhle et al., 2020).

According to Prigogine (1980), to fully grasp the concept of self-organization in social systems, one must understand the critical transitional changes between order and chaos, stability and confusion. In stable states, the system works like it used to work. However, this state of "*equilibrium*" also means that there is little room for new or unexpected developments. The system needs to be pushed to the "chaos zone" to achieve new developments and results. If the system always operates in a stable equilibrium, through its familiar practices, innovative developments do not emerge. Characteristics that promote stability inhibit creativity because they only allow group members to do things that are guided by the same frames of mental models as usual (Salazar, 2002). The space between stability and instability—the edge of chaos—is where higher-level learning and creativity take place.

Ståhle (1998) analysis of Prigogine's research is used in this study. Key characteristics (i.e., requirements) for all self-organizing systems: *entropy*, *state of far-from-equilibrium,* and *momentums of bifurcation* are extracted from Prigogine's work by Ståhle. The focus of this study is the self-organization of knowledge within a group.

Next, the process and requirements of self-organization in social systems are clarified.

Entropy, a fundamental concept in thermodynamics, also plays a crucial role in the self-organization process. In the context of social systems, it refers to the information that a system produces to generate iteration but also increased disorder and randomness in itself. The systemic basis of self-organization is interaction, a rich exchange of information among its components. The more that the system exchanges information, i.e., communicates, the more the level of entropy increases. However, there are certain requirements. First, the communication dynamics in the system must be non-linear, and second, it must include both positive and negative responses. Both are needed for fruitful discrepancies and confusion to emerge. Positive feedback alone creates unanimity and negative feedback alone prevents continuity (Ståhle, 1998). When communication increases and the dynamics of communication include positive and negative forms, the process is called iterative: a cyclic feedback process that is continuous and sensitive, which allows the information produced by a system to be quickly transmitted throughout the whole system. Increased entropy also means information that cannot be utilized. Contrary to earlier beliefs, Prigogine saw high entropy levels not as a waste, but instead as a necessary component of self-organization. According to Ståhle (1998), Prigogine argued that a self-organizing system always produces waste and abundant information. Paradoxically, this uselessness is also necessary for a system's evolution.

In Prigogine's theory, increasing entropy levels are also associated with information chaos and disorganized, unclassified, or unappreciated knowledge. Entropy also brings uncertainty, imbalance, and confusion into the system through increased communication and different perspectives. It is important to note that high entropy means greater disorder, wasted resources, lost information, and uncertainty in the system. For a social system, this means abundant communication and production of ideas, and different angles of information without any certainty as to whether they will prove useful. Entropy is key in the self-organization process. Without increased entropy levels, the system (group) stays in a stable state (equilibrium) where new developments can only be small steps without any radical transformations.

To sum up, the entropy level of the system is based on its iterative quality, i.e., the *frequency* of information exchanged and the number of *different viewpoints*, as well as the balance between *positive and negative feedback* and *equal participation (power balance)* of the system components.

2.3. From comfort equilibrium to chaotic far-from-equilibrium

As described above, high entropy levels are necessary for a social system to move from its stable state, *equilibrium*, toward self-organization. When the system moves from its comfort zone to "*far-from-equilibrium*", it handles increasing entropy levels, uncertainty, and confusion. If the system can handle the disharmony and confusion of increased entropy levels, it has a chance to produce order out of chaos; thus, it is crucial to avoid making interpretations and crystallizing information too early in the system, as this can hinder the system's ability to reach the needed "chaos zone." Far-from-equilibrium might lead to the creation of something genuinely new—a new order, new knowledge, out of chaos.

2.4. Bifurcation: a zone between determinism and free choice

The new order in a system's self-organization includes momentums of bifurcation. "In principle, a bifurcation is simply the appearance of a new solution" (Prigogine, 1980, p. 105). Bifurcation always produces a change that is not a logical continuation of the previous structure (Prigogine, 1980, p. 105), and thus bifurcation as an event is always also a source of innovation (Prigogine and Nicolis, 1989, p. 74). The change of the system to the new equilibrium state happens suddenly. At the point of bifurcation, the system rejects a large amount of information, causing the amount of entropy to decrease and a new order to emerge. Bifurcation requires chaos or a state of far-from-equilibrium, as stated by Prigogine (1980, p. 105), Prigogine and Stengers (1984, p. 169), Prigogine and Nicolis (1989, p. 74), and Ståhle (1998).

According to Ståhle (1998) analyses of Prigogine, bifurcation refers to a phenomenon characterized by irreversible changes. As noted by Prigogine and Nicolis (1989), bifurcation is a catalyst for innovation and diversification, leading to the emergence of new solutions in the system. This transition occurs abruptly, akin to a sudden leap, which the term "crystallization" accurately portrays. At the bifurcation moment, the system relinquishes a significant amount of information, resulting in a decrease in entropy and the creation of a new order.

Bifurcation moments are critical in comprehending the irreversible changes that occur in self-organization. These systems are pushed beyond their initial equilibrium states through fluctuations, leading them to the bifurcation moments where multiple new options can be established. At this point, the system must choose between the available alternatives, and upon passing through the bifurcation moment, the system assumes a new configuration with new properties and structures.

In this study, our conceptual and analytical approach to the social system's self-organization can be seen to bring forth interesting emphases to modern learning theories. Following this, the related learning research context is presented, and concepts from our analytical frame are interconnected.

2.5. Reflection on the learning research context

In this study, the knowledge co-construction process is seen to include affordances for the group to self-organize and support higher-level learning, such as critical thinking, problem-solving, creativity, and reflection, and a more comprehensive understanding of the subject matter. The social system's self-organization is an iterative process with the unpredictable resonance between its components. Similarly, in socio-constructivism, learning is an ongoing, iterative process built from unscripted dialogue, interaction, the agency of participants, and artifacts (Lonka, 2015; Lonka et al., 2018). Learning goes beyond individual cognitive processes and also includes distributed, group, and social aspects (Hontvedt et al., 2023). Learners engage in cycles of inquiry, reflection, and action during the construction process, accommodating existing knowledge structures and refining their understanding. Material facets and the role of artifacts in collaborative processes are emphasized (Papert and Harel, 1991; Paavola and Hakkarainen, 2005; Stahl, 2006; Stahl and Hakkarainen, 2020).

How is this kind of iterative co-construction process supported in practice? Lonka and Ahola (1995) summarized three general principles for student-activating teaching and learning methods in higher education: first, starting with activating, and diagnosing the previous knowledge and understanding of the participants (e.g., brainstorming and generating ideas); second, supporting the learning process and making the learning processes overt to the discussion in various ways (e.g., guided discussions and editing shared artifacts); and third, providing both formative and summative assessments throughout the learning process (e.g., learning diaries, evaluation discussions, and further editions of artifacts). These three phases of learning processes are iterative and cyclical, where a longer learning cycle (such as an entire course) consists of shorter cycles of tutorials, lessons, or group discussions. Project-, inquiry-, and problem-based learning are examples of methods that activate students (Barron et al., 1998; Bereiter, 2002; Hmelo-Silver, 2004; Pedaste et al., 2015). In more ordinary settings, such as student-activating lectures, student-activating principles are also effective (e.g., McKeachie, 1999; Lonka and Ketonen, 2012). Lonka (2012) and Lonka et al. (2018) has presented a synthetic *Engaging Learning Model*, adding engagement and interest to the cyclic learning process described above: starting by catching interest and curiosity, then maintaining interest, and finally, deepening the interest of the students during the cycle (based on Hidi and Renninger, 2006). Activating students' ideas and potential misconceptions may also trigger confusion when previously held ideas are challenged. In all these phases, tutoring, and scaffolding of learning are important (e.g., Muukkonen et al., 2005).

In this study, the system-theoretic approach of self-organization and analytical tools is used to bring an understanding of how technology, as an active component of the system, adds value to the co-construction process. Entropy refers to the iterative information exchange that a system produces, including increased randomness and uncertainty in a system. In a learning context, this relates to the concept of surprise, which is one epistemic emotion. Epistemic emotions, which relate to knowledge and the generation of knowledge, are critical drivers of cognitive performance and engagement in learning (Vogl et al., 2020). As fundamental epistemic emotions, surprise, interest, confusion, and curiosity are linked with antecedents (e.g., cognitive dissonance) and outcomes (e.g., knowledge creation) and are therefore critically important for learning. Created confusion (imbalance) can turn toward discoveries, and curiosity promotes new levels of thinking. The findings of several studies imply surprise focuses attention, enhances memory, triggers interest and curiosity, and indirectly influences motivation. Sudden changes can push the system beyond its boundaries, sparking epistemic emotions (Renninger and Hidi, 2015; Noordewier et al., 2016; Vogl et al., 2020).

Increased entropy nudges a system from equilibrium toward chaos, a vital state for self-organization. Similarly, learning also requires tolerance of negative emotions such as confusion, boredom, and frustration, as they serve to stretch our understanding (Lonka et al., 2018). Rather than signs of failure, these challenging emotions are a natural part of the learning process. Specifically, confusion can stimulate constructive learning and deep understanding (Craig et al., 2004; D'Mello and Graesser, 2014). The benefits of confusion for learning depend on how it arises within tasks and how students manage it (Lehman et al., 2012; Lodge et al., 2018; Arguel et al., 2019). Effective confusion resolution is crucial for successful learning outcomes, as unresolved confusion can dampen interest in learning (D'Mello and Graesser, 2012). Furthermore, the learning environment should facilitate confusion management through timely feedback and align cognitive disequilibrium with task context for problem-solving (D'Mello and Graesser, 2014). However, it is important to note that individual differences significantly affect how students experience confusion in the learning process.

These kinds of entropy-driven dynamics, triggering bursts of epistemic emotions, can catalyze bifurcation moments, innovations, and new solutions in the system. In collaborative learning, bifurcation momentums in self-organization can be harnessed toward the concept of collaborative emergence (Sawyer and DeZutter, 2009), which refers to the emergence of a new product or creative outcome, a collective chance for something novel and appropriate to occur. Cognitive processes, distributed across participants and artifacts, contribute to collaborative emergence (Sawyer, 2006, 2009). Characterizing this unrestricted process is a free-flowing collaboration marked by equal participation and flexible actions. Additionally, spontaneous responses to changing situations empower the occurrence of something novel (Sawyer, 2006, 2009).

3. Materials and methods

3.1. Case study

3.1.1. Context and participants

The research context of this case study was a science teacher education course. The focus was on pedagogical content knowledge, especially how to include science education research knowledge of students' conceptual understanding while planning science instruction. The course was part of a 1-year program of pedagogical studies, required in Finland to achieve formal teacher qualification in addition to master-level studies in the teaching subject, e.g., mathematics or physics. The program was conducted in English. In the course, there were two small groups, with a total of eight participating student teachers of mathematics or science. Due to the international study program, the students' backgrounds were notably diverse, including their experience with educational technology. Their ages ranged from 20 to 57 years. Participants' backgrounds varied, ranging from a career-changer transitioning from information technology to teaching, to a 4^{th}-year student with little experience beyond their university studies. All participants gave their written consent to the planned research.

Focus is placed on one lesson (90 min) in this study. The collaborative activities of the lesson aimed at supporting reflection on *why it is important for a teacher to track students' preconceptions* and *what pedagogical aspects to take into consideration when planning science and engineering practices for physics lessons*. Group 1 (G1) comprised one female and three male participants, and group 2 (G2), three female and one male participants. There was one physics major in each group. The whole group tested the software during an earlier lesson on the previous day, so the functionality of the software was not totally new to them.

3.1.2. AI-directed speech recognition technology for self-organization of knowledge

The core concept of the AI-directed speech recognition technology used in this case study was based on theories of collaborative learning and knowledge co-construction. The main idea is that participants can focus on the flow of conversation and produce "notes", digital artifacts, while speaking, without the additional effort required by typing or writing.

In this research, technology is seen as an inseparable part of the system. Technology provides a novel means for collaboration by enabling the co-creation of digital artifacts (word clouds and collective notes) through spoken contributions. The artifacts produced by technology during co-construction activities are seen as active components of the system, as a part of a group of learners. The notion of artifacts has been used interchangeably in learning sciences (Damşa, 2014). Artifacts are instruments to mediate learners' actions in dialogue and problem-solving, as well as engaging them in knowledge construction (Säljo, 1999). "Externalized and materialized artifact" refers to instruments that promote the evolution of understanding and guide personal or collective inquiry further (Ritella and Hakkarainen, 2012). The role of material artifacts during the construction process has more recently shifted from artifacts solely as learning outcomes to artifacts as an active part of learning. Artifacts have the agency to create situations where ideas emerge from a process of interaction with them (artifacts). "The pedagogical message throughout is that the student should welcome the artifact's agency as a positive force that drives the creative process forward" (Sawyer, 2022).

In this study, the AI-directed speech recognition technology used combines new disruptive technologies: (1) speech recognition as a novel collective way to produce artifacts orally without typing; and (2) artificial intelligence to transform recorded oral thoughts into visual, materially embodied digital artifacts. Automatic Speech Recognition (ASR) is a computerized procedure that decodes and transcribes spoken language, typically converting it into written text. Many empirical investigations have provided substantial evidence to support the positive impact of ASR on language learning in particular (Jiang et al., 2023). Speech recognition is one of the most complex areas of computer science—involving linguistics, mathematics, and statistics. The vagaries of human speech have made its development challenging. Numerous factors can impact word error rates, such as pronunciation, accent, pitch, volume, and background noise. Reaching human parity—meaning an error rate on par with that of two humans speaking—has long been the goal of speech recognition systems.

The core characteristics of the software used in this study are as follows. It is (1) built to work in the English language. After recording an oral group discussion, (2) the software materializes transcription, based on group discussion as a digital, frequency-based word cloud. It is (3) possible to move each word in the cloud separately, (4) resize the words, and (5) add new words and phrases to the word cloud by typing on a keyboard.

3.1.3. Pedagogical planning and co-construction activities

The pedagogical planning of this 90-min lesson was framed to apply the *Engaging Learning Model* (ELM) (Lonka, 2012; Lonka et al., 2018) with a simplified short learning cycle:

1. Phase 1—to organize activities and set goals, share current understanding, capture interest, and generate ideas.
2. Phase 2—to facilitate inquiries through a shared artifact and engage participants in knowledge co-construction process.
3. Phase 3—to assess learning gains and engage the participants in deepening their interest and motivating future learning.

    1. Phase 1: The aim of this phase was to start the co-construction process: that is, to activate previous understanding and to start brainstorming based on oral discussions. The setting was a small group learning discussion around the table. The large format touch device was not in use during this phase. The participants were able to write their own notes. Microphones were set up on the table and voice-driven AI software recorded the learning discussion in the background.
    2. Phase 2: Co-construction activity during the second phase aimed at deepening the learning process. The group activity consisted of editing the digital word cloud and organizing the single artifacts from the cloud on a large format touch device. Phase 2 began when the digital word cloud, externalized from oral group discussions and transcriptions, became visible, making digital artifacts available for knowledge co-construction.
    3. Phase 3: In Phase 3, the co-construction activity of the groups was to present and share their learning outcomes for evaluation. Edited word clouds were available and displayed on the large format touch device during the groups' presentations. Table 1 summarizes the co-construction phases, activities, and technology setup.

Table 1 Co-construction phases, activities, and technology setup in each phase.

Lesson's phases (Length of each phase)	Co-construction activities	Technology setup
1. Phase (25 min) to activate previous understanding and to start brainstorming based on oral discussions	Small group discussion around the table	Software records group discussions and creates transcription of small group discussions. Large-format touch device is not in use.
2. Phase (35 min) aimed at deepening the learning process	Digital word cloud artifact is visible and the words in it are editable for the group on a large-format touch device.	From transcription, the software produces, based on frequency, a digital word cloud and editable artifacts (words) in it. Word cloud is displayed on a large-format touch device.
3. Phase (30 min) present and share learning insights	Group presentations to the whole class and the teacher.	The edited word clouds of both groups are available on the large-format touch device during the group's own presentation

3.2. Data collection and analysis

The video allowed us to examine the role of technology in group processes in depth (Derry et al., 2010). The present study employed the following data from the larger data corpus (240 min in total). Field notes were also collected. First, the video data were scrutinized. The main data set consisted of a 90-min lesson (co-construction Phases 1, 2, and 3). The data were collected using a video camera focused on a group with a microphone placed on the table in two separate classrooms. The selected video data (150 min in total) formed an entity where technology was used throughout the entire co-construction process. The data consisted of videos from both small groups' co-construction Phases 1 and 2, and from Phase 3 when the whole group and a teacher were together in one classroom.

The analysis proceeded as follows. The selected video recordings (150 min) were transcribed verbatim. Next, the video and transcripts were analyzed. The video data and transcriptions were imported into the Atlas.ti software. The data processing and analysis using Atlas.ti was carried out by the first author. The focus of the analysis was identified by the involvement of co-authors.

Table 2 Qualitative subsection criteria of entropy.

Qualitative subsections of the concept of entropy	Definition	Example of the thematic speech episode	Example from the data
(1) Reinforces the previous speech acts	Contains a positive confirmation or concurring opinion for the previous speech turns	Agreed that starting with a pre-task involving questions can help identify students' preconceptions.	"Yeah, yeah, actually, I was thinking about that too…ike, it might be good to have like some kind of… not like an exam or anything, but just some kind of questions to find out what the preconceptions are."
		Editing the word cloud	"Yeah, that's true. Let's put that in the middle."
(2) New insight to the discussion	Participant presents new insight or viewpoint to the conversation	Mentioned that the exercise of interviewing someone without a physics background is a useful starting point for understanding preconceptions	"The point was to get somebody who didn't have the background in physics to think about it and try to get to the intuitive preconceptions that they might have."
		Introducing ideas on how to teach abstract physics phenomenon	"We use experiments to convey abstract ideas?"
(3) Opposite viewpoint	Participant disagrees or brings the opposite or different point to the discussion.	Expressing doubt about including difficulties or problems in the word cloud summary.	"Those are sort of generic, so I don't know." "Come on. No, I mean like how is everything useful in physics. We don't need it."
		Working with artifacts and discussing. Selecting artifacts from the word clouds.	"We have to change all of this."
(4) Other	A speech act which cannot be categorized under any ofthe three entropy categories above.	Moving artifacts	"Just making space for these."
		Sharing emotional state on Phase 1	"I'm so tired."
		Off-topic artifacts	"Weekend, Sunday."

Definitions, examples of the thematic speech episodes, and examples from the data.

Initially, the participants' verbal acts were analyzed using the code-and-count technique. Each participant's discussion turns in both groups was divided into speech episodes, constituting the unit of analysis (Linell, 1998, 2009). In this study, a speech episode is defined as a thematically meaningful unit of interactional exchange. A new episode begins when the discussion turns to the next participant. By applying the code and count technique and calculating frequencies of speech episodes in co-construction processes, initial analyses of the changes in entropy levels (amount of information exchanged) during three different co-construction phases in both groups were performed.

Next, the speech episodes were analyzed according to the analytical framework (Ståhle, 1998). Our focus in the analysis was directed toward the identification of the concepts of (1) entropy, (2) far-from-equilibrium, and (3) bifurcation momentums across the three phases of co-construction.

Following the counting of speech episodes, the concept of *entropy* was further analyzed using qualitative content analysis (Braun and Clarke, 2006; Strijbos et al., 2006). Quantified speech episodes were also analyzed qualitatively, to understand the content dimensions and entropy subsections of the speech episodes. The unit of analysis was a speech episode that could be thematically categorized to represent characteristics of entropy. Three key entropy dimensions of variation in speech episodes were identified. In this study, it meant that the speech episodes either: (1) reinforced the previous speech act, (2) brought a new insight to discussion, or (3) represented an opposite viewpoint. Category (4) "Other" was created for some of the speech episodes that were not relevant to any of these three categories. Subsection criteria of entropy, examples of the thematic analysis, and data are presented in Table 2.

In the third phase of analysis, units that captured a meaningful unity from the pre-service teachers' verbal and non-verbal actions, characterizing signs of the system's behavior in *far-from-equilibrium* state, were identified. In this study, this state was indicated by signs of confusion, frustration, or challenges manifested in the group members' verbal acts or behavior. Video-based interaction analysis (Jordan and Henderson, 1995) consists of the in-depth microanalysis of how people interact with one another, their physical environment, and the documents, artifacts, and technologies in that environment. The transcripts were segmented into topical episodes based on the substantive content of the speech and non-verbal behavior. Simultaneously occurring episodes were considered to be separate episodes. The talk and actions in these episodes were analyzed in terms of signs of confusion, frustration, or a sense of challenge. These signs were also systematically compared with entropy levels (number of speech episodes) and the existence and utilization of digital artifacts.

Finally, the analysis of speech episodes was continued through the concept of *bifurcation* in social systems' self-organization using qualitative content analysis. Bifurcation moments in this research are referred to as crystallized learning insights resulting from co-construction. The speech episodes from the co-construction Phase 3 were analyzed and compared to the edited word cloud artifacts.

The general focus was on the complex arrangement of verbal, visual, and material conduct through which the participants interacted during co-construction, including the role of the digital artifacts in the process (cf. Wohlwend, 2009; Theobald, 2012).

4. Results

4.1. Description of co-construction phases

The aim of co-construction Phase 1 was to begin the co-construction process with small group discussion: To activate previous understanding and to start discussing "*How to take into consideration physics abstract situations when planning teaching, so that students can understand what the key aspects of the motion and force as a physics phenomenon are.*" Technology was an "invisible" component in Phase 1. Speech-recognition software recorded the learning discussion in the background. Participants were able to write their own notes. Both groups' discussions were content related. All participants in both groups were able to contribute something new to the discussion, which followed the principle of alternation.

Phase 2 aimed at deepening the co-construction process. This phase began when both groups, in their own classrooms, saw the production of the recording, the word cloud artifact, for the first time. The aim was to produce a summary of pedagogical ideas for presentations by using the artifacts available. From a technical point of view, AI-directed speech recognition technology merged and transacted individuals' speech, and externalized spoken discussion using data mining functions based on the recorded group discussions. Technology formulated the digital word cloud from Phase 1 group discussions. Each individual's intangible shared thoughts were brought and merged. In Phase 2, technology became visible to the groups, so it also became one component of the system. Both groups went in front of the touch device and started to edit the artifacts while continuing their discussions (Figure 1). Selected artifacts from the word clouds in both groups represented the learning task at hand. The classified and organized artifacts represent content knowledge, and examples of these from group 1 are "Physics," "Experiment," and "Force," and from group 2 are "Motion," "Laws," and "Preconceptions."

In Phase 3, the groups presented and shared their learning insights. In this third phase of co-construction, the two groups came together with a teacher and presented their findings, one group at a time. During the presentations, all members from both groups had the opportunity to participate in sharing ideas and thoughts from the previous phases.

4.2. The role of AI-directed speech recognition technology on co-construction in terms of self-organization of knowledge

To answer the research question "*What is the role of the AI-directed speech recognition technology in the co-construction process in terms of self-organization of knowledge?,*" the findings in terms of entropy are first illuminated. *In the initial analysis, each participant's interactional speech episodes during the three co-construction phases were quantified. This was done with the intention of interpreting variation in entropy levels across both groups.* The code and count approach, as explained in the data analysis, was used to get a full view of the interactional turns, which indicates the entropy levels in social systems' self-organization processes. The results showed that participants' discussion turns, or "entropy levels," increased rapidly in Phase 2, when the digital word cloud became visible for the groups to edit on the large format touch device. Tables 3, 4 show the frequencies of the discussion turns of every participant in groups 1 and 2 in co-construction Phases 1 and 2. The groups worked on these two phases in their own classrooms. Table 5 shows the frequency of the discussion turns of every participant and the teacher in co-construction Phase 3 when both groups and the teacher came together in the same classroom. Entropy production was highest in Phase 2, concerning every participant in both groups (ID01–ID04 in G1 and ID05–ID08 in G2). The frequency of speech episodes was notably higher compared to Phases 1 and 3. The increased discussion turns mean more exchange of information and interaction between participants: Group 1: Phase 1 (variation between) 8–12 turns, Phase 2: 19–29, and Phase 3: 4–12. Group 2: Phase 1: 7–9 turns, Phase 2: 18–34, and Phase 3: 7–12.

A self-organizing system must have the ability to produce and accumulate entropy. Speech episodes of every participant in both groups were analyzed to determine how speech episodes were distributed among the group members and thus reveal the power balance of communication. The results demonstrate that every participant in each group contributed to the discussion evenly and no one dominated the discussions excessively. As presented earlier, the self-organization of the system requires the production and accumulation of entropy (the number of speech episodes), but it is also critical that all information is considered equal.

As noted above, producing and increasing entropy levels (the amount of communication) is one aspect of entropy, equally valued information is second, and the qualitative contents and meanings of speech episodes is third. Next, the content of the speech episodes was analyzed to understand the meanings and qualitative dimensions in terms of entropy in self-organization. It was shown by our results that speech episodes represented various dimensions of entropy: (1) reinforcement of previous speech episodes, (2) new viewpoints, and (3) opposite viewpoints. In the data, there were also many speech episodes that did not fit into any of these three categories. The (4) "Other" section includes speech episodes that do not represent dimensions of entropy. This section represents off-topic or random episodes. Tables 3, 4 show that in Phase 1, participants in both groups generated discussion by reinforcing previous speech episodes or presenting new viewpoints to the discussion. There was only one opposite viewpoint in each group in Phase 1. These results in Phase 1 reflect that co-construction activity is represented by a typical small group learning discussion where everyone shares their opinions on the topic at hand. Speech episodes in such activities might not contain that much dissonance: the focus is more on sharing one's own thoughts, one participant at a time. However, disharmony is needed for the system to be able to self-organize. In Phase 2, when the digital artifacts were available for groups to edit, there were more opposing viewpoints and also more new viewpoints from content knowledge that participants had brought to the discussion when compared to the results from Phase 1.

In summary, the results of the first analysis phase showed that both groups increased entropy levels by exchanging more information during Phase 2 when digital artifacts were visible and available to edit. The distribution of speech episodes among participants reflected equal participation. Speech episodes during

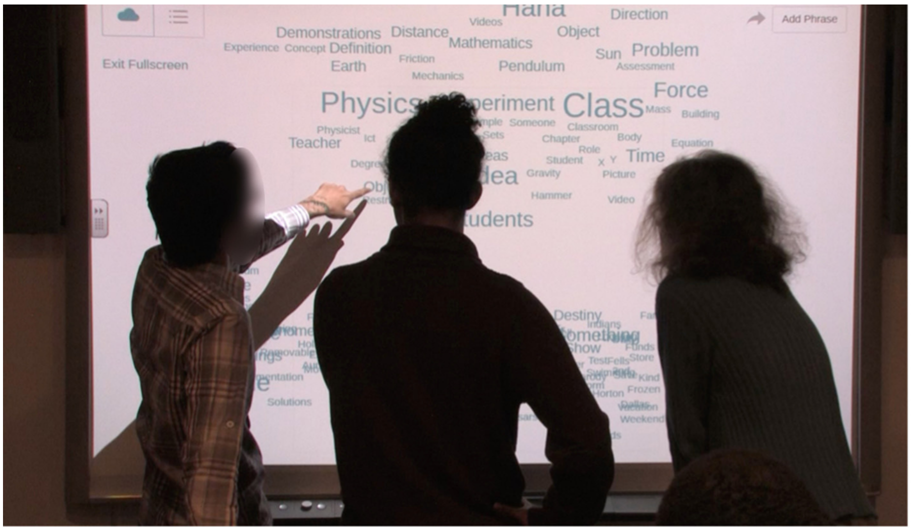

FIGURE 1
Group 1 is editing the word cloud artifact in Phase 2.

Table 3 Discussion turns and characteristics of speech episodes of group 1 in Phase 1 and Phase 2.

Participant	Total speech episodes	Reinforces previous speech episodes	New viewpoints	Opposite points	Other
Phase 1–group 1					
ID05	9	3	3	0	3
ID06	12	5	3	1	3
ID07	9	3	2	0	4
ID08	8	1	1	0	6
Phase 2–group 1					
ID05	29	12	8	3	8
ID06	20	9	5	5	5
ID07	21	8	7	2	6
ID08	19	7	8	1	4

Table 4 Discussion turns and characteristics of speech episodes of group 2 in Phase 1 and Phase 2.

Participant	Total speech episodes	Reinforces previous speech episodes	New viewpoints	Opposite points	Other
Phase 1–group 2					
ID05	9	4	3	0	2
ID06	7	4	2	1	0
ID07	9	4	3	0	2
ID08	9	4	3	0	3
Phase 2–group 2					
ID05	26	5	10	3	8
ID06	34	8	15	5	6
ID07	24	6	10	2	6
ID08	18	3	6	1	4

Table 5 Discussion turns and characteristics of speech episodes of group 1, group 2, and the teacher (everyone in the same classroom) in Phase 3.

Participant	Total speech episodes	Reinforces previous speech episodes	New viewpoints	Opposite points	Questions	Other
Phase 3–group 1, group 2, teacher together						
ID01	12	2	2	0	0	3
ID02	6	3	3	0	0	0
ID03	5	2	2	0	1	0
ID04	4	0	3	0	1	0
ID05	12	2	5	0	0	5
ID06	7	2	3	0	1	3
ID07	12	4	2	0	1	5
ID08	10	2	3	1	0	4
Teacher	7	2	3	0	2	3

this phase included key qualitative characteristics of entropy, not only reinforcing the previous opinions but also bringing opposed or new views of content knowledge into the process. Based on the results, the role of the artifacts—materialized digital word clouds—represented a new source of information, which increased entropy production and accumulation of entropy (amount of discussion turns), and also triggered new and opposing points into the discussion. Presence and working with the artifacts brought out the iterative nature of the co-construction process, which is explained next.

4.2.1. From equilibrium to far-from-equilibrium

As presented above, small group learning discussion represented a well-known and familiar educational practice where a group of learners share their thoughts. From the system's self-organizing perspective, during this activity, both groups were near equilibrium during Phase 1, based on earlier results regarding entropy levels and the qualitative content of speech episodes. Also, results from the interaction analysis showed that in Phase 1, there were no signs of confusion, uncertainty, or disorder which are indicators of a far-from-equilibrium state in our analysis.

Far-from-equilibrium is a central state in the system's self-organization. The far-from-equilibrium state opens the possibility to enter the "chaos zone", which is imperative for the self-organization of knowledge. Far-from-equilibrium or chaos zone in social systems does not necessarily mean a real chaotic atmosphere, but there must be a redundancy of contradictory and opposed ideas, which are signs of uncertainty and confusion.

Selected examples of the data, which indicate a far-from-equilibrium state in the system's self-organization, are presented next. The results show that when Phase 2 began and digital word clouds became visible to the groups on the large format touch device for the first time, signs of surprise and confusion appeared. Excerpt 1 demonstrates the first signs of perplexity when group 1 saw the word cloud for the first time in their own classroom. Mistakes in speech recognition led to the appearance of unexpected and confounding words within the word cloud.

Based on the results of interaction analysis, the digital word cloud artifact played the key role and pushed the groups gently from the stable *equilibrium* state, which, based on the results, refers to the small group discussion in Phase1, to the *far-from-equilibrium* phase, which also evoked epistemic emotions among group members (Figure 2).

Excerpt 1 Group 1 sees the digital word cloud for the first time.

Transcription	Non-verbal actions on video	Notes related to artifacts
ID04: Whoa!	Everyone in their own seats looking at the word cloud.	Word cloud is visible to the group for the first time.
ID01: Destiny. No one said that.	ID01, ID02, ID03, ID04 are laughing. Looking at each other.	
ID03: over. First time (speaking on top of each other)		
ID02: hammer		
ID04: rooftop	ID04 is laughing.	
ID02: Gravity		
ID03: There's a haha!	ID01, ID02, ID03, ID04 are laughing.	
ID03: Should we sort out?		
ID01: Yes, just sort out		

Signs of surprise, confusion, and frustration, which indicate a far-from-equilibrium state in this study, were the focus of our analysis. Laughing is one of the signs of raised epistemic emotions, representing surprise, and confusion. The occurrence of laughter implied that the situation held unexpected and therefore puzzling elements. The appearance of a word cloud revealed conversation points that the participants had not mentioned in the earlier discussion during Phase 1. Furthermore, laughter was also prompted by mistakes made by the technology. These errors relaxed the atmosphere and provided material to deepen the co-construction process.

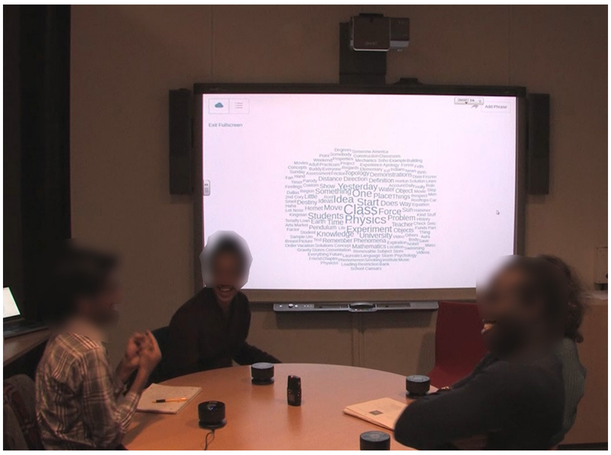

FIGURE 2
Group 1 when they saw the word cloud artifact for the first time.

Excerpt 2 Group 2 started to work with the word cloud.

Transcription	Non-verbal actions on video	Notes related to artifacts
ID08: Yeah, ok. Recording lots of disturbing words.	The group moves from their seats to the front of the screen.	Word cloud is visible to the group for the first time.
Move those to the side and then put the others in?	Every group member started to move the words of the cloud on the screen at the same time.	
ID05: and rearrange.		
ID07: Too many fingers.		
ID08: Okay, maybe one or two on.	One of the participants went to sit down because of the confusion caused by everyone editing the cloud at the same time (Later the participant came back in front of the screen).	

Excerpt 3 Group 1 is editing the word cloud.

Transcription	Non-verbal actions on video	Notes related to artifacts
ID04 Soho.. That's interesting?	Pointing out the word "Soho" on the screen.	Soho, water, haha, direction, demonstration, distance are all words in the word cloud.
Water Did you say water probably… who knows…		
ID01 But yeah, okay, so let's see. haha, direction, demonstration, distance. Okay, so how do we collect all this?	Moving the single artifacts on the screen.	
ID04 Okay should we now try to work on this?... I think what we should do... I think (ID04)		

Excerpt 4 Group 2 is editing the word cloud.

Transcription	Non-verbal actions on video	Notes related to artifacts
ID08 This is strange. Nobody talked about this?	Scratching the head, looking at the screen.	Everyone is in front of the screen.
ID06 Yeah?	Looking at the screen.	
ID05 Integrity we already have. Oh, we have it somewhere here.	Looking for the right word from the word cloud, pointing on the cloud.	
ID07 Intuitive we have but not? integrity		

Meanwhile, when group 2 saw their artifact for the first time in their classroom, they immediately went in front of the screen and started to work with the concepts of the word cloud. There were also signs of confusion among these participants (Excerpt 2) regarding the words in the word cloud and also challenges when they started to work with the cloud and move the single words on the screen.

The word cloud artifacts themselves included surprising and confusing elements. Wrong off-topic recognitions in the word clouds triggered expressions of epistemic emotions. Confusion and challenges are needed to move groups toward self-organization. The word clouds did not show the full sentences or the logical order of the discussion. The word cloud result was surprising, a mix of key concepts from participants' speech and also sound material from incorrect speech recognition by the technology. In normal open-ended conversation, vagaries of human speech and unexpected sounds occur, which were also captured by the recordings. The technology "took in" every sound from the other sound sources (such as sneezes, coughing, background noise, or talking over each other) and processed these other sounds as if they were speech. Speakers' accents can also cause occurrences that increased the error rate of speech recognition. Both types of wrong recognitions caused words that were not discussed to appear in the word clouds, which caused observable confusion when the editing of the word clouds continued. Excerpt 3 shows this while group 1 was editing the cloud.

Also in group 2, the group members were confused by the wrongly recognized words in the word cloud (Excerpt 4). They tried to find the concepts from the word cloud that they had been talking about during Phase 1.

Editing the word cloud created frustration among the group members. Group 2 was trying to change the size of the words to make them more visible. After changing the size of one word, the word cloud rearranged itself and mixed the already selected and arranged concepts into different places on the screen (Excerpt 5).

Despite these challenges, the groups did not give up on the task, and they continued to engage by discussing and shaping the word clouds to represent their thoughts regarding pedagogical aspects of physics teaching.

As stated above, in the process of self-organization, generating entropy (exchange of information), accumulating entropy (different viewpoints), confusion, and tolerance of challenges are

Excerpt 5 Group 2 is editing the word cloud.

	Non-verbal actions on video	Notes related to artifacts
ID06: "How do you change the size…"	Looking at others	Trying to change the size of the artifact. The word cloud was rearranging already moved artifacts after the group had added words to it.
ID07: "Double click."		
ID05: "Oooh."	Surprised expression	
ID06: "Oh no?."	Sad expression	
ID06: "It never works when we want it to work?."		
ID07: "Yeah?."		

crucial elements in reaching a far-from-equilibrium state. It is critical that the system should also be able to dissipate entropy to self-organize. In the context of social systems, this means that information needs to be analyzed and evaluated. The editing of the word cloud and discussions of the shared artifacts represents the analysis and evaluation of the information gained. In this study, technology, as one system component, scaffolded co-construction, and at the same time, the self-organization of knowledge by squeezing and visualizing group conversation into a word cloud format. The word cloud summarized parts of the exchanged information from Phase 1. With the visualized artifact, it was possible to grasp the thoughts of previous Phase 1 conversations, despite the wrong recognized words in the cloud, and dive more collectively into evaluating the topic at hand.

It is critical to avoid making interpretations and the final crystallization of information too early. If the system can handle the confused situation, a turmoil of chaos may produce order out of chaos. This makes it possible for new structures or innovations to arise. When information is analyzed, valued, or modeled, the entropy and chaos with uncertainty decrease. Working with digital artifacts led the groups to spend more time discussing the learning task at hand again. Crystallization of information was not done immediately after the Phase 1 conversations. In Phase 2, the co-construction process included a vivid process of classifying and organizing the information scaffolded by artifacts on the screens. Classifying of artifacts and discussions focused on building shared understanding and deepening perceptions, which started during Phase 1. From the systemic self-organizing view, when participants discussed, organized, and classified the artifacts into a summary of their pedagogical ideas, the atmosphere of confusion and uncertainty started to lift as the work with the clouds and discussions continued.

4.2.2. Momentums of bifurcation points

The phenomenon of self-organization in a social system includes characteristics of seemingly spontaneous emergence of new structural features and order following a period of disruption, "far-from-equilibrium". As presented in the results above, in this study, digital word clouds were the key elements in leading and maintaining groups at a far-from-equilibrium state and also for keeping the systems and the process iterative. The word cloud artifacts also generated opposing viewpoints among the learners, which is essential for creating the conditions for bifurcations. As the iteration continues and the system operates long enough in this state by classifying information, the system can become increasingly receptive, and move closer to the *bifurcation moments* that can result in a new state, a new structure, or the crystallization of knowledge. Bifurcation is a catalyst for innovation and diversification, leading to the emergence of new solutions in the system.

A content analysis was conducted on the speech episodes from Phase 3 where both groups came together and presented their findings from the two earlier co-construction phases. First, in this study, the materialized word cloud itself can be seen to represent a significant turning point in the co-construction process, which can be referred to as a bifurcation point. The concepts in the clouds were based on groups' discussions so the clouds can be seen to be a crystallization of the exchanged information, which the technology made visible as one active component of the system. The technological artifact transformed the groups' monological flow of thought into collective views, visually merging consecutive speech turns from Phase 1 into a word cloud. The common visual output offered a turning point for the group's co-construction process. Misinterpretations in the clouds served as surprising elements and reflective mirrors that increased communication and confusion, thereby stimulating learners to collaboratively construct knowledge.

It may be too daring to talk about real bifurcation moments in this study, which in their wider meaning can be seen to lead to significant shifts in learners' understanding and knowledge structures in the context of the social systems. As a result, instead of using the concept of bifurcation moment, further editing the digital word cloud could be seen to represent a "*bifurcation space*" (Figure 3), where there are many different branches to work with knowledge and where groups are allowed to make free choices by picking up, removing, and adding the artifacts they like. The choices made in this state depend on the system and cannot be predicted in advance. These acts, every decision that the groups made regarding artifacts they selected from the word cloud, can be seen to crystallize information further and represent *learning by insights,* which are more likely to emerge in this kind of short co-construction cycle than bifurcation moments in their wider meaning. Interacting with and organizing digital artifacts, the system abandons a large amount of information and thus the amount of entropy decreases. It can be said that learning insights were reached by both groups during this moment of the process. Based on content analysis, bifurcation moments in this study represented collective and crystallized thoughts about the learning insights of the co-construction cycle. Excerpt 6 presents examples of learning insights from group 2 while they presented their thoughts in Phase 3.

5. Discussion

The purpose of this study was to analyze knowledge co-construction as a self-organization process and the role of

FIGURE 3
On the **left**, word cloud after Phase 1. On the **right**, edited word cloud after Phase 2. Both from group 2.

Excerpt 6 Group 2 is sharing learning insights in Phase 3.

ID05: First, we thought based on yesterday's example, that it is better to clear up the preconceptions.
ID06: Yeah, so to find out what the preconceptions are… with some kind of task
ID08: And after we do the pre-task, then we can identify the existing knowledge that the students bring to the class.
ID06: Like first find out what they think will happen, then show them a video of what's actually happened?. They can like see that okay, I was right or I was wrong.
ID06: Then you can start talking about like, what happens in the phenomenon. We came up with the idea of teaching them about the different Newtonian laws, because usually they are not too clear to the students.
ID08: And then it's important while teaching that there are so many ways to teach, so give many different examples and different contexts.
ID07: So the overall idea is to have a plan, to find differences between preconceptions, and Newton's laws. And, the solution involves sort of showing videos, working in groups, repeating various examples. And then coming and discussing meanings within the group and then afterwards with the teacher about what kind of effects you could see… to point out the differences between the expectations of Newton's laws, and then the preconceptions that the students had.

technology as its catalyst. Novel AI-directed speech recognition technology and the artifacts it generates were deployed to scaffold the knowledge co-construction process of pre-service teachers in two groups in a science education context. Analysis was conducted through key characteristics of the social system's self-organization based on Ståhle (1998). By taking a systems theory approach, a pattern was found that pointed out the key elements in enabling a group to use technology as a catalyst to self-organize in knowledge co-construction.

The self-organization of knowledge involves a sequence of iterative steps. First, the system must produce entropy, promoting a rich exchange of information and increased randomness. Next, an increase in entropy leads the system from its stableness, "equilibrium", toward a far-from-equilibrium state, in which confusion and uncertainty are necessary elements for knowledge to be analyzed, evaluated, and finally crystallized. When a system can operate with increased entropy, classifying and reflecting on information and tolerating challenges, it has the opportunity to evolve and innovate a new order through bifurcation moments.

In this study, a system referred to an entity composed of a student group, technology, artifacts produced by students and technology, and co-construction activities. Based on the results of this study, the role of technological artifacts to enable signs of self-organization during knowledge co-construction was significant in both groups. First, the results showed that in every participant's discussion turns, the "entropy levels of the systems" increased rapidly in Phase 2, when technology became an active component of the system. The digital word cloud and artifacts became visible for the groups to edit on large-format screens. The number of speech episodes—entropy—increased in both groups. Also, the speech episodes included various qualitative forms of entropy, such as opposing views and new ideas. Visible and perplexing artifacts and engaging collaboration with them pushed the groups from stable small-group discussion in an equilibrium state (Phase 1) toward far-from-equilibrium (Phase 2). Without the digital word cloud artifacts, mandatory signs of self-organization (increasing entropy levels, different forms of entropy, and far-from-equilibrium states) were not found in either group through data analysis. The presence of technology and artifacts spawned more communication, and participation was distributed equally among the group members. Technology, as an inseparable part of the system, brought surprising and confusing elements to co-construction. It can be said that the technological artifacts destabilized a familiar educational structure, small-group discussion, and gave a new boost to the process, enabling observable signs of self-organization processes. The word clouds and artifacts can be seen to present a sort of *bifurcation moment*. The participants' earlier discussions were crystallized and materialized by technology. AI-directed technology offered bifurcation, a concrete and visible summary draft that scaffolded groups to construct knowledge and gain learning insights into the topic at hand. Working with the digital artifacts represented a *bifurcation space* where groups were able to follow different branches and co-constructed knowledge further.

The results of this study raise questions: how can new technologies be built and used to support novel kinds of co-construction processes and the emergence of beneficial epistemic emotions? What new practices then emerge? The prevailing assumption has often been that advanced educational technology should adjust the learning process in real-time to suit the learners' current knowledge and skills. This adjustment aims to prevent uncertainty or confusion. Simply put, common wisdom holds that confusion should be avoided during learning and rapidly resolved if and when it arises (D'Mello et al., 2014). But when it comes to the construction process, epistemic emotions such as surprise or confusion relate to learning itself and have an object focus on knowledge construction (Pekrun et al., 2017). Positive learning experiences should include enough challenges and even confusion (Vilhunen et al., 2023). If technology-mediated processes are scripted and built to keep learners "always on track," predictability is maximized, and emergence is minimized. This might diminish the appearance of epistemic emotions, which in the co-construction process have the potential, based on our results, to lead the group toward deeper collaborative efforts and self-organization. As D'Mello et al. (2014) claim, confusion plays an important role during complex learning activities and, if appropriately regulated, it can cause learners to process the material more deeply to resolve their confusion. Learning environments need to challenge learners substantially to elicit critical thought and deep inquiry. To continue, Scardamalia and Bereiter (2014) state that scripted instruction stands in contrast to the emergent character of human action, creative thinking, and constructive work with ideas. To foster higher-level "emergent," technologies can play a role as enablers, not only supporting productive interaction between people but also promoting engagement between people and ideas, which can lead to the emergence of innovations (Scardamalia and Bereiter, 2014).

Teachers play an important role in fostering learning and emotional scaffolding (Halonen et al., 2016; Vilhunen et al., 2023). The role of a modern teacher is also to plan and facilitate learners to move beyond their comfort zones, transitioning from stable equilibrium states toward dynamic far-from-equilibrium situations. In such situations, learners have opportunities and freedom to interact and develop innovations together. According to the findings of this study, if co-construction activities and the use of technology are creative and framed pedagogically meaningfully, learners acquire opportunities for deeper knowledge construction processes. Thus, as pedagogical experts, teachers can employ elements of surprise and confusion in relation to technology-mediated learning activities. Embracing mistakes, incorporating randomness, and confronting challenges are essential aspects of innovative knowledge co-construction. The three-phase, student-activating pedagogical framework (Lonka et al., 2018), also used in this study, could aid in the planning of technology-enhanced learning activities that may induce confusion, unexpected turns, or other challenges designed to evoke epistemic emotions.

In the current era of expanding artificial intelligence (AI), it is necessary for existing pedagogical structures to harness technologies more extensively to support self-organizational elements in the learning context. Instead of merely speeding up the tasks and giving ready-made answers, EdTech could be a catalyst for co-construction and offer surprising collective twists to evoke epistemic emotions that feed creativity, engagement, and higher-level learning elements as a natural part of non-linear pedagogy and collaborative learning. When technology scaffolds self-organizational processes and fosters the emergence of knowledge co-construction, teachers can focus more on crucial aspects of human learning that also contribute to a communal and resilient future such as facilitating students' social-emotional wellbeing, school engagement, and metacognitive skills.

The significance of this study emerges from a systemic approach where reflections regarding the role of technology in co-construction were illuminated by our results. It was demonstrated that technology, as an active component of the system, can catalyze co-construction. The analysis and results of this study underscore the need for systemic thinking in a Computer-Supported Collaborative Learning (CSCL) context. Emerging technologies and digital artifacts are more than just interactors, they are increasingly active agents within the system. Together with the group of learners, they can facilitate the whole system toward the self-organization of knowledge, thus creating a valuable space for innovative learning as knowledge creation.

The limitations of this research are acknowledged. It is recognized that the results cannot be generalized due to the case study nature of this research. Follow-up studies are necessitated in other learning contexts and on other learning elements, as well as the use of different educational technologies. Despite these limitations, this study creates interesting openings for applying the analytical tools of self-organization in technology-mediated co-construction.

Data availability statement

The datasets presented in this article are not readily available as the research rights are only applicable to this project. Requests to access the datasets should be directed to NH, niina.halonen@helsinki.fi.

Author contributions

NH: writer and owner of the manuscript, planning from the start of the research project, involvement in every phase, and has done most of the writing and analysis. PS: planning, writing, and commenting the manuscript. KJ: planning and involvement of the co-construction activities of the study and involvement in analysis. SP: writing and commenting the manuscript. KL: head of the whole research project, planning and commenting of the manuscript, and involvement in analysis. All authors contributed to the article and approved the submitted version.

Funding

NH received a personal research grant from the Jenny and Antti Wihuri Foundation. This work was also supported by the

Digiconsumers Strategic Research Council (SRC) at the Academy of Finland (1352545).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Arguel, A., Lockyer, L., Kennedy, G., Lodge, J. M., and Pachman, M. (2019). Seeking optimal confusion: a review on epistemic emotion management in interactive digital learning environments. *Inter. Lear. Environ.* 27, 200–210. doi: 10.1080/10494820.2018.1457544

Barron, C. (2003). Problem-solving and EAP: themes and issues in a collaborative teaching venture. *English Specific Purp.* 22, 297–314. doi: 10.1016/S0889-4906(02)00016-9

Barron, O. E., Kim, O., Lim, S. C., and Stevens, D. E. (1998). Using analysts' forecasts to measure properties of analysts' information environment. *Account. Rev.* 73, 421–433.

Bereiter, C. (2002). Design research for sustained innovation. *Cogn. Stud.* 9, 321–327. doi: 10.11225/jcss.9.321

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., and Krathwohl, D. R. (1956). *Handbook I: Cognitive Domain.* New York: David McKay.

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qualit. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Brod, G., Kucirkova, N., Shepherd, J., Jolles, D., and Molenaar, I. (2023). Agency in educational technology: Interdisciplinary perspectives and implications for learning design. *Educ. Psychol. Rev.* 35, 25. doi: 10.1007/s10648-023-09749-x

Craig, S., Graesser, A., Sullins, J., and Gholson, B. (2004). Affect and learning: an exploratory look into the role of affect in learning with AutoTutor. *J. Educ. Media* 29, 241–250. doi: 10.1080/1358165042000283101

Damşa, C. I. (2014). The multi-layered nature of small-group learning: Productive interactions in object-oriented collaboration. *Int. J. Comput. Supp. Collabor. Lear.* 9, 247–281. doi: 10.1007/s11412-014-9193-8

Derry, S. J., Pea, R. D., Barron, B., Engle, R. A., Erickson, F., Goldman, R., et al. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *J. Learn. Sci.* 19, 3–53. doi: 10.1080/10508400903452884

Dillenbourg, P., Järvelä, S., and Fischer, F. (2009). "The evolution of research on computer-supported collaborative learning," in *Technology-Enhanced Learning* (Springer, Dordrecht) 3–19. doi: 10.1007/978-1-4020-9827-7_1

D'Mello, S., and Graesser, A. (2012). Dynamics of affective states during complex learning. *Lear. Instr.* 22, 145–157. doi: 10.1016/j.learninstruc.2011.10.001

D'Mello, S., and Graesser, A. (2014). Confusion and its dynamics during device comprehension with breakdown scenarios. *Acta Psychol.* 151, 106–116. doi: 10.1016/j.actpsy.2014.06.005

D'Mello, S., Lehman, B., Pekrun, R., and Graesser, A. (2014). Confusion can be beneficial for learning. *Lear. Instr.* 29, 153–170. doi: 10.1016/j.learninstruc.2012.05.003

Haataja, E. S., Tolvanen, A., Vilppu, H., Kallio, M., Peltonen, J., and Metsäpelto, R. L. (2023). Measuring higher-order cognitive skills with multiple choice questions–potentials and pitfalls of Finnish teacher education entrance. *Teach. Teacher Educ.* 122, 103943. doi: 10.1016/j.tate.2022.103943

Halonen, N., Hietajärvi, L., Lonka, K., and Salmela-Aro, K. (2016). Sixth graders' use of technologies in *learning*, technology attitudes and school well-being. *Eur. J. Soc. Behav. Sci.* 18, 2307–2324. doi: 10.15405/ejsbs.205

Härkki, T., Vartiainen, H., Seitamaa-Hakkarainen, P., and Hakkarainen, K. (2021). Co-teaching in non-linear projects: A contextualised model of co-teaching to support educational change. *Teach. Teacher Educ.* 97, 103188. doi: 10.1016/j.tate.2020.103188

Hennessy, S., and Murphy, P. (1999). The potential for collaborative problem solving in design and technology. *Int. J. Technol. Design Educ.* 9, 1–36. doi: 10.1023/A:1008855526312

Hidi, S., and Renninger, K. A. (2006). The four-phase model of interest development. *Educ. Psychol.* 41, 111–127. doi: 10.1207/s15326985ep4102_4

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educ. Psychol. Rev.* 16, 235–266. doi: 10.1023/B:EDPR.0000034022.16470.f3

Hontvedt, M., Prøitz, T. S., and Silseth, K. (2023). Collaborative display of competence: A case study of process-oriented video-based assessment in schools. *Teach. Teacher Educ.* 121, 103948. doi: 10.1016/j.tate.2022.103948

Jiang, M. Y. C., Jong, M. S. Y., Lau, W. W. F., Chai, C. S., and Wu, N. (2023). Exploring the effects of automatic speech recognition technology on oral accuracy and fluency in a flipped classroom. *J. Comput. Assist. Learn.* 39, 125–140. doi: 10.1111/jcal.12732

Jordan, B., and Henderson, A. (1995). Interaction analysis: Foundations and practice. *J. Lear. Sci.* 4, 39–103. doi: 10.1207/s15327809jls0401_2

Kangas, K., Seitamaa-Hakkarainen, P., and Hakkarainen, K. (2013). Design thinking in elementary students' collaborative lamp designing process. *J. Design Technol. Educ.* 18, 30–43.

Lehman, B., D'Mello, S., and Graesser, A. (2012). Confusion and complex learning during interactions with computer learning environments. *Internet. Higher Educ.* 15, 184–194. doi: 10.1016/j.iheduc.2012.01.002

Linell, P. (1998). *Approaching Dialogue: Talk, Interaction and Contexts in Dialogical Perspectives (Vol. 3).* Amsterdam: John Benjamins Publishing. doi: 10.1075/impact.3

Linell, P. (2009). *Rethinking Language, Mind, and World Dialogically.* New York, NY: Information Age Publishing.

List, C., and Pettit, P. (2011). *Group Agency: The Possibility, Design, and Status of Corporate Agents.* Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199591565.001.0001

Lodge, J. M., Kennedy, G., Lockyer, L., Arguel, A., and Pachman, M. (2018). "Understanding difficulties and resulting confusion in learning: An integrative review," in *Frontiers in Education* (Frontiers Media SA) 49. doi: 10.3389/feduc.2018.00049

Lonka, K. (2012). "Engaging Learning Environments for the Future: The 2012 Elizabeth W. Stone Lecture," in *The road to information literacy: Librarians as facilitators of learning: IFLA Publications Series 157,* eds. R., Gwyer, R., Stubbings, G., Walton (Berlin/Munich: De Gruyter Saur) 15–30. doi: 10.1515/9783110281002.15

Lonka, K. (2015). *Oivaltava Oppiminen.* Helsinki: Otava.

Lonka, K., and Ahola, K. (1995). Activating instruction: How to foster study and thinking skills in higher education. *Eur. J. Psychol. Educ.* 10, 351–368. doi: 10.1007/BF03172926

Lonka, K., and Ketonen, E. (2012). How to make a lecture course an engaging learning experience? *Stud. Learn. Soc.* 2, 63–74. doi: 10.2478/v10240-012-0006-1

Lonka, K., Makkonen, J., Berg, M., Talvio, M., Maksniemi, E., Kruskopf, M., et al. (2018). *Phenomenal learning from Finland.* Helsinki: Edita.

McKeachie, W. J. (1999). "Teaching, learning, and thinking about teaching and learning," in *Higher Education: Handbook of Theory and Research* (Dordrecht: Springer Netherlands) 1–38. doi: 10.1007/978-94-011-3955-7_1

Mitchell, M. (2009). *Complexity: A Guided Tour.* Oxford: Oxford University Press.

Muukkonen, H., Lakkala, M., and Hakkarainen, K. (2005). Technology-mediation and tutoring: How do they shape progressive inquiry discourse? *J. Learn. Sci.* 14, 527–565. doi: 10.1207/s15327809jls1404_3

Noordewier, M. K., Topolinski, S., and Van Dijk, E. (2016). The temporal dynamics of surprise. *Soc. Person. Psychol. Compass* 10, 136–149. doi: 10.1111/spc3.12242

Paavola, S., and Hakkarainen, K. (2005). The knowledge creation metaphor–An emergent epistemological approach to learning. *Sci. Educ.* 14, 535–557. doi: 10.1007/s11191-004-5157-0

Paavola, S., and Hakkarainen, K. (2009). "From meaning making to joint construction of knowledge practices and artefacts–A trialogical approach to CSCL," in *Computer Support Collaborative Learning Practices: CSCL2009 Conference Proceedings.*

Papert, S., and Harel, I. (1991). Situating constructionism. *Constructionism* 36, 1–11.

Pedaste, M., Mäeots, M., Siiman, L. A., De Jong, T., Van Riesen, S. A., Kamp, E. T., et al. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educ. Res. Rev.* 14, 47–61. doi: 10.1016/j.edurev.2015.02.003

Pekrun, R., Vogl, E., Muis, K. R., and Sinatra, G. M. (2017). Measuring emotions during epistemic activities: The epistemically-related emotion scales. *Cogn. Emot.* 31, 1268–1276. doi: 10.1080/02699931.2016.1204989

Prigogine, I. (1976). "Order through fluctuations: Self-organization and social systems," in *Evolution and consciousness: Human systems in transition,* eds. E. Jantsch and C. H. Waddington (Boston: Addison-Wesley) 93–133.

Prigogine, I. (1980). *From Being to Becoming: Time and Complexity in the Physical Sciences*. San Francisco: W.H. Freeman.

Prigogine, I., and Nicolis, G. (1989). *Exploring Complexity: An Introduction.* New York: Freeman.

Prigogine, I., and Stengers, I. (1984). *Order Out of Chaos: Man' s Dialogue With Nature*. New York: Bantam.

Renninger, K. A., and Hidi, S. (2015). *The Power of Interest for Motivation and Engagement*. London: Routledge. doi: 10.4324/9781315771045

Ritella, G., and Hakkarainen, K. (2012). Instrumental genesis in technology-mediated learning: From double stimulation to expansive knowledge practices. *Int. J. Comput. Support. Collabor. Learn.* 7, 239–258. doi: 10.1007/s11412-012-9144-1

Salazar, A. J. (2002). "Self-organizing and complexity perspectives of group creativity: Implications for group communication," in *New Directions in Group Communication,* ed. L. W. Frey (London: Sage) 179–200. doi: 10.4135/9781412990042.n10

Säljo, R. (1999). Learning with interactive graphical representations. *Learn. Instr.* 9, 303–425.

Sanina, A., Kutergina, E., and Balashov, A. (2020). The Co-Creative approach to digital simulation games in social science education. *Comput. Educ.* 149, 103813. doi: 10.1016/j.compedu.2020.103813

Sawyer, K. (2009). "Writing as a collaborative act," in *The psychology of creative writing* 166–179. doi: 10.1017/CBO9780511627101.012

Sawyer, R. K. (2006). Educating for innovation. *Think. Skills Creat.* 1, 41–48. doi: 10.1016/j.tsc.2005.08.001

Sawyer, R. K. (2022). The dialogue of creativity: Teaching the creative process by animating student work as a collaborating creative agent. *Cogn. Instr.* 40, 459–487. doi: 10.1080/07370008.2021.1958219

Sawyer, R. K., and DeZutter, S. (2009). Distributed creativity: How collective creations emerge from collaboration. *Psychol. Aesthet. Creat. Arts* 3, 81. doi: 10.1037/a0013282

Scardamalia, M., and Bereiter, C. (2014). Smart technology for self-organizing processes. *Smart Lear. Environ.* 1, 1–13. doi: 10.1186/s40561-014-0001-8

Scheel, L., Vladova, G., and Ullrich, A. (2022). The influence of digital competences, self-organization, and independent learning abilities on students' acceptance of digital learning. *Int. J. Educ. Technol. High. Educ.* 19, 1–33. doi: 10.1186/s41239-022-00350-w

Schleicher, A. (2018). *World Class*. Paris: OECD Publishing. doi: 10.1787/9789264300002-en

Ståhle, P. (1998). *Supporting a system's capacity for self-renewal (Research Report 190)*. Department of Teacher Education, University of Helsinki.

Ståhle, P. (2008). "The dynamics of self-renewal: A systems-thinking to understanding organizational challenges in dynamic environments," in *Organisational Capital* (London: Routledge) 133–160. doi: 10.4324/9780203885215.ch7

Ståhle, P., Mononen, L., Tynjälä, P., and Kallio, E. K. (2020). "Systems thinking and adult cognitive development," in *Development of Adult Thinking* (London: Routledge) 191–207. doi: 10.4324/9781315187464-12

Stahl, G. (2006). *Group Cognition: Computer Support for Building Collaborative Knowledge (Acting with Technology).* New York, NY: The MIT Press. doi: 10.7551/mitpress/3372.001.0001

Stahl, G., and Hakkarainen, K. (2020). *Theories of CSCL. International Handbook of Computer-Supported Collaborative Learning.* New York, NY: Springer. doi: 10.1007/978-3-030-65291-3_2

Stenalt, M. H. (2021). Researching student agency in digital education as if the social aspects matter: students' experience of participatory dimensions of online peer assessment. *Assess. Eval. Higher Educ.* 46, 644–658. doi: 10.1080/02602938.2020.1798355

Strijbos, J. W., Martens, R. L., Prins, F. J., and Jochems, W. M. (2006). Content analysis: What are they talking about? *Comput. Educ.* 46, 29–48. doi: 10.1016/j.compedu.2005.04.002

Theobald, M. (2012). Video-stimulated accounts: Young children accounting for interactional matters in front of peers. *J. Early Childhood Res.* 10, 32–50. doi: 10.1177/1476718X11402445

Vilhunen, E., Chiu, M. H., Salmela-Aro, K., Lavonen, J., and Juuti, K. (2023). Epistemic emotions and observations are intertwined in scientific sensemaking: A study among upper secondary physics students. *Int. J. Sci. Math. Educ.* 21, 1545–1566. doi: 10.1007/s10763-022-10310-5

Vogl, E., Pekrun, R., Murayama, K., and Loderer, K. (2020). Surprised–curious–confused: Epistemic emotions and knowledge exploration. *Emotion* 20, 625. doi: 10.1037/emo0000578

Wohlwend, K. E. (2009). Mediated discourse analysis: Researching young children's non-verbal interactions as social practice. *J. Early Childh. Res.* 7, 228–243. doi: 10.1177/1476718X09336950

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 14 February 2024
DOI 10.3389/feduc.2024.1295824

OPEN ACCESS

EDITED BY
Wang-Kin Chiu,
The Hong Kong Polytechnic University, China

REVIEWED BY
Ririn Ambarini,
Universitas PGRI Semarang, Indonesia
Hideyuki Kanematsu,
Suzuka College, Japan
Lingyun Sun,
Zhejiang University, China

*CORRESPONDENCE
Zhiyong Fu
✉ fuzhiyong@tsinghua.edu.cn
Min Liu
✉ liumin@bnu.edu.cn

RECEIVED 17 September 2023
ACCEPTED 02 February 2024
PUBLISHED 14 February 2024

CITATION
Liu W, Zhu Y, Li Y, Fu Z, Sun Y, Hong X, Li Y and Liu M (2024) Co-making the future: judges' insights on transdisciplinary creativity and global collaboration in the China-U.S. young maker competition. *Front. Educ.* 9:1295824. doi: 10.3389/feduc.2024.1295824

Co-making the future: judges' insights on transdisciplinary creativity and global collaboration in the China-U.S. young maker competition

Wei Liu[1], Yancong Zhu[1], Yin Li[2], Zhiyong Fu[2]*, Yuanbo Sun[3], Xinhui Hong[4], Yanru Li[5] and Min Liu[6]*

[1]Beijing Key Laboratory of Applied Experimental Psychology, National Demonstration Center for Experimental Psychology Education (Beijing Normal University), Faculty of Psychology, Beijing Normal University, Beijing, China, [2]Academy of Arts and Design, Tsinghua University, Beijing, China, [3]School of Design and Arts, Beijing Institute of Technology, Beijing, China, [4]Xiamen Academy of Arts and Design, Fuzhou University, Xiamen, China, [5]Office of International Exchange and Cooperation, Beijing Normal University, Beijing, China, [6]Institute of Comparative and International Education, Beijing Normal University, Beijing, China

This paper examines the future of maker education through an analysis of feedback from judges in the China-U.S. Young Maker Competition. Drawing on inputs from 36 judges from diverse backgrounds in academia, industry, and sponsoring companies, the study uses thematic analysis of interviews, feedback, and focus group discussions to uncover key educational trends. It highlights critical themes such as transdisciplinary creativity, real-world application, sustainability, cross-cultural collaboration, and innovation mindset. The research reveals a trend towards integrating various academic fields to boost creative problem-solving and application in real-life scenarios. Sustainability is identified as a crucial component, pointing to the need for environmentally aware education. The study also emphasizes the importance of cross-cultural collaboration for global interconnectedness and adaptive problem-solving, alongside fostering a continuous innovation mindset in students. Concluding with future directions for maker education, the paper advocates for an experiential, inclusive, and forward-looking educational approach. It underscores the importance of a broad curriculum that integrates entrepreneurial skills, promotes lifelong learning, and enhances global connectivity. This study provides insights for educators, policymakers, and practitioners, offering a streamlined roadmap for advancing maker education in a rapidly evolving global context.

KEYWORDS

maker education, thematic analysis, cross-cultural perspectives, transdisciplinary creativity, China-U.S. collaboration

1 Introduction

In our current era, marked by a wave of open innovation often described as "mass entrepreneurship and innovation," we are seeing significant changes in our society. This change is fueled by a burst of creative thinking and problem-solving (Basadur and Hausdorf, 1996; Asheim et al., 2007; Chatterji et al., 2014; Clapp et al., 2016; Hepp, 2020). At the heart of this change is the "maker" movement, a blend of open innovation principles that has created a lively and interactive community. This movement is about two main things: a strong love for technology and a commitment to making innovative ideas come to life (Martinez and Stager, 2013; Halverson and Sheridan, 2014; Kolb et al., 2014; Lindtner, 2015). The maker movement, which started in the do-it-yourself (DIY) and hacker cultures of Europe and America, has now become well-known worldwide. It is known for encouraging innovation, sharing openly, being involved hands-on, and always looking to improve the quality of life. Events like the Maker Faire have become symbols of this movement, attracting support from both governments and communities (Tabarés and Boni, 2023). In China, the DIY culture has been popular since the 1980s and has grown to include activities like making custom furniture and assembling personal computers, showing a dedication to creative and practical work (Lazonick, 2004; Williamson, 2016; Wen et al., 2022).

In this setting, the China-U.S. Young Maker Competition stands out as an important event. It is more than just a competition; it is a place where creative minds come together to solve big global problems, like environmental sustainability and climate change, with their inventive ideas. It is a mix of different ideas and cultures and has been a place for fostering innovation and developing talent for over a decade.

This paper takes a close look at this period of vibrant innovation and the maker culture. Instead of focusing on the participants, as many studies do, we turn our attention to the judges of the China-U.S. Young Maker Competition. This new focus is intended to give us a better understanding of how the maker movement affects education and has a wider impact on society. We use a qualitative thematic analysis to look closely at the data collected from the judges of the competition. This approach lets us explore their viewpoints in depth, giving us a better understanding of what happens in the competition. This method is different from most of the research done before, which usually focuses on what the participants experience. With this study, we aim to answer an important question: How do the judges' views in the China-U.S. Young Maker Competition show and shape the current trends in maker education and its overall effect on society? This question is at the heart of our study, as we look to uncover deeper insights into how innovation, education, and cultural exchange interact in this unique international competition.

2 Related works

2.1 Deepening creativity and innovation in maker education

The maker movement's integration into educational systems represents a transformative shift, redefining learning by nurturing creativity and innovation (Weng et al., 2022). This movement, emerging from DIY and hacker cultures, challenges the traditional educational framework by promoting active, hands-on learning and creative problem-solving. It enables students to evolve from passive learners to active creators, integrating diverse disciplines from arts, science, and technology (Dym et al., 2005). This transdisciplinary approach not only enhances creative thinking but also fosters essential problem-solving skills, addressing the needs of today's rapidly changing world.

Educational competitions within the maker movement serve as powerful catalysts for this transformation (Liu et al., 2021). They provide real-world challenges that inspire students to apply their theoretical knowledge in practical contexts, thereby promoting a culture of inventive thinking and collaborative problem-solving. These competitions are more than just contests; they are platforms where students can showcase their technical skills and creative prowess. They motivate students to break free from conventional thought processes and explore innovative possibilities. This encourages the cultivation of a generation of innovators who are not only technically proficient but also creatively confident.

The maker movement, particularly through these competitions, plays a crucial role in shaping future innovators. These events challenge students to create and innovate, pushing them to develop solutions that are both imaginative and technically sound (Miettinen, 2000). They inspire a spirit of exploration and discovery, essential for fostering a culture of continuous learning and adaptation. As students engage in these competitions, they learn to navigate complex problems, work collaboratively, and think critically, preparing them for the challenges of the modern world (Martin, 2015; Chakraborty et al., 2023; Tablatin et al., 2023). The maker movement encourages inclusivity and diversity in problem-solving approaches. By bringing together students with varied backgrounds and skill sets, it fosters an environment where different perspectives are valued and explored. This diversity is critical in driving innovation, as it leads to a richer pool of ideas and solutions. As such, the maker movement and its associated competitions are pivotal in developing well-rounded individuals who are equipped to contribute to and thrive in a world that values creativity and innovation.

2.2 Enhancing a comprehensive approach in educational competitions

Incorporating Human-Centered Design (HCD), Human-Computer Interaction (HCI), and User eXperience (UX) design within maker education exemplifies a multidimensional approach that focuses on fostering innovation, empathy, and responsibility (Ren et al., 2019; Al Mahmud and Soysa, 2020; Yang et al., 2022; Liu et al., 2023; Zhu et al., 2024). HCD in maker education extends beyond the creation of functional solutions; it's about crafting projects that are impactful and prioritize human needs and experiences. This perspective encourages students to think from the end-user's viewpoint, leading to designs that are not only effective but also empathetic and meaningful. HCI and UX design play a crucial role in this educational paradigm. These disciplines ensure that technology is not merely technically advanced but also accessible and engaging. They prompt students to consider how users interact with technology, emphasizing the importance of intuitive design and meaningful user experiences. Such an approach is essential in preparing students to develop technology that is not just functional but enjoyable and efficient to use. The integration of these principles in maker education and competitions represents a shift towards a more holistic view of technological development. It encourages students to create solutions that consider the broader context of their use, including accessibility, usability, and practicality. This approach is vital in nurturing a generation of innovators who are adept at balancing technical proficiency with thoughtful, user-centered design.

In educational competitions, this comprehensive approach influences how projects are developed and evaluated. The focus extends beyond technical skill to encompass how well projects align with principles of HCD and technological intuitiveness. It instills in students a deeper understanding of the importance of creating solutions that are not only innovative but also considerate of the users' needs and experiences (Desmet et al., 2023; Lachheb et al., 2023). Ultimately, this approach in maker education cultivates a sense of responsibility among students towards creating more inclusive and user-friendly technology. It prepares them to become creators who are not only technically skilled but also mindful of the human aspect of technological innovation. They learn to create solutions that are not just effective but also enrich users' lives, setting a new standard for how technology is designed and utilized.

2.3 Expanding the role of the competition in global innovation and sustainability

Since its inception in 2014, the China-U.S. Young Maker Competition has grown into a significant platform for fostering cross-cultural innovation and collaboration. With its impressive participation—over 50,000 individuals contributing to more than 14,000 projects—it has become a major driver of innovation in numerous cities and universities in both countries. This competition is not only a testament to the creativity and technical skills of its participants but also a reflection of the growing importance of global collaboration in education and innovation.

The competition's alignment with the United Nations Sustainable Development Goals (UN SDGs) underscores its commitment to contributing to global challenges through creative and sustainable innovation (Carlsen and Bruggemann, 2022; Lafont-Torio et al., 2024). By incorporating these goals, the competition encourages participants to develop projects that are not only technologically advanced and creatively rich but also address important issues like climate change, sustainable urban development, and responsible consumption. This focus on the UN SDGs elevates the competition from being merely a technical showcase to a platform for meaningful global impact.

The diverse range of projects that emerge from this competition highlights the potential of young innovators to contribute to sustainable solutions for the world's most pressing problems. From addressing environmental concerns to promoting social equity, the projects align with various SDGs, showcasing the competition's role in driving forward these crucial global agendas. The competition serves as an important model for how educational initiatives can integrate creativity, technology, and sustainability. It demonstrates the value of fostering a mindset among young innovators that prioritizes not just technical proficiency but also a deep understanding of the broader societal and environmental implications of their creations.

3 Methodology and data analysis

3.1 Approach to qualitative inquiry

A qualitative research methodology was employed to explore the perspectives of judges in the competition (Sanders and Stappers, 2012; de Bont, 2021). This approach was chosen for its effectiveness in capturing detailed insights into judges' experiences, decision-making processes, and evaluative criteria. The qualitative method allowed for an in-depth exploration of judges' viewpoints on creativity, innovation, and the criteria they applied within the competition. The flexibility inherent in qualitative research enabled adjustments and refinements in our approach as new themes and insights emerged, ensuring a dynamic and comprehensive inquiry process.

3.2 Engagement and data gathering

In this research, our focus was on data collection and analysis, distinctly separate from the roles of judges or mentors within the competition. By adopting a non-participatory, analytical stance, we were able to ensure an objective approach to data collection, which was essential for accurately capturing and interpreting the judges' perspectives. This methodology allowed us to gather data while maintaining the integrity and authenticity of the judges' experiences and viewpoints.

The data collection process was meticulously organized to encompass a broad spectrum of perspectives from the judges, who hailed from varied professional backgrounds. The judging panel consisted of 36 individuals: 24 from the academic sector, 8 from industry, and 4 representing the sponsor companies of the competition. Their areas of expertise covered a wide range, including computer sciences, creative industry, entrepreneurship, and industrial design, thus offering a rich and comprehensive collection of professional insights. Additionally, the judges' average age of 45.18 years brought together a mix of seasoned experience and contemporary perspectives, further enriching the data collected for our analysis.

- Interviews: in-depth interviews with the judges, conducted in both structured and semi-structured formats, were instrumental in gathering detailed insights into their assessment criteria, the challenges they encountered, and their viewpoints on the projects evaluated. The varied professional backgrounds of the judges, encompassing areas like computer sciences and the creative industry, offered a multifaceted understanding of the competition's evaluation process.
- Judging criteria and feedback reviews: an extensive analysis of the judging criteria and written feedback from the judges resulted in over 50 pages of detailed notes and reflections. This part of the data collection was key in deciphering the various criteria and considerations used by judges from different fields such as entrepreneurship and industrial design in their evaluations.
- Focus group discussions: organizing focus group discussions with the judges allowed for an in-depth exploration of their collective experiences and viewpoints. These discussions, enriched by the judges' diverse professional backgrounds, provided deeper insights into their consensus and differing opinions. We collected extensive transcripts from these focus groups, totaling over 30,000 words. This substantial dataset offered a thorough understanding of the judges' collective thought processes, decision-making, and the dynamics of their evaluations.

3.3 Analysis of qualitative data

The analysis process involved a structured thematic approach to describe the collected data (Braun and Clarke, 2023):

- Initial coding and organization: the data from interviews, feedback reviews, and focus group discussions were initially categorized into broad thematic areas to facilitate organization and analysis.
- Refined theme development: further examination of these initial codes led to the identification of refined themes, revealing deeper insights into judges' perspectives on creativity, innovation, and their evaluative processes within the competition.
- Narrative construction and synthesis: the final stage involved constructing a coherent narrative that integrated these themes, providing a comprehensive portrayal of the judges' roles and impacts on the competition.

This methodology and analysis approach provided a thorough and detailed understanding of the judges' perspectives in the competition. The qualitative analysis was pivotal in uncovering significant insights into the judges' contributions to fostering an environment of cross-cultural innovation and understanding in this notable educational event.

4 Findings and discussion

The thematic analysis represented in Table 1 serves as the bedrock for the findings detailed in sections 4.1 through 4.6. Initiated with an exhaustive examination of various data sources, including structured interviews with judges, feedback forms, and focus group discussions, each source contributed indispensable insights integral to comprehensively understanding the impact of the maker competition. For instance, the theme 'Embracing Transdisciplinary Creativity' discussed in section 4.1 was principally derived from the data obtained through judges' interviews. These interviews provided rich qualitative insights, particularly highlighting the judges' appreciation for the integration of different disciplines in project development. The feedback forms complemented this theme, offering concrete examples of the judges' focus on cross-disciplinary skills and artistic expression. In similar fashion, sections 4.2 through 4.6 explore themes such as 'Prioritizing Practical Application and Real-World Impact' and 'Nurturing Global Awareness Through Sustainability'. These themes were significantly informed by focus group discussions, which revealed depth in perspectives concerning the importance of real-world applications, sustainability practices, and the value of cross-cultural collaboration in the projects.

The analytical approach in this study extended beyond simple data gathering. It involved a thorough process of organizing and interpreting the data to ensure it was relevant and well-supported by strong evidence. This careful process was crucial in creating a narrative that is both engaging and solidly based on empirical data. By clearly showing the sources of our data and including a wide range of perspectives, we sought to strengthen the credibility and trustworthiness of our findings. The resulting narrative is not only complete but also reflects the comprehensive and varied nature of our data analysis. The table does more than just list themes; it demonstrates the detailed analytical process we undertook. It illustrates how each level of coding, from the first to the third, is linked to specific data sources, thus providing a thorough view of the judges' perspectives. This organized approach ensures that our findings are supported by a wide array of data, from direct quotes in interviews to shared insights from focus group discussions. Therefore, the table serves as evidence of the thoroughness and depth of our analysis, highlighting the careful thought and scrutiny that support the conclusions of our study.

4.1 Embracing transdisciplinary creativity

In the rapidly evolving educational sector, the judges' emphasis on transdisciplinary creativity in the competition is a guiding light for future educational trends. This approach underscores the necessity of integrating various disciplines, such as arts, science, and technology, to cultivate a more holistic understanding and application of knowledge. This blend enriches students' learning experiences, equipping them with a broader skill set and fostering a mindset that transcends conventional academic boundaries. A judge eloquently stated, "Blending disciplines in maker projects leads to more comprehensive and creative solutions, bridging the gap between theory and practical application." This philosophy underscores the imperative to prepare students for the complexities of the modern world, where problems often require multifaceted solutions that draw on a range of disciplines. By embracing this transdisciplinary approach, maker education can become a powerful tool for nurturing versatile, innovative thinkers capable of addressing contemporary challenges with creativity and depth.

- Integrate diverse disciplines: advocate for the inclusion of diverse subjects in maker education, promoting projects that combine arts, science, and technology.
- Foster creative problem-solving: encourage educational programs that emphasize creative thinking and innovative problem-solving approaches.
- Nurture versatile skill sets: develop curriculum structures that build versatile skills, preparing students for multidisciplinary challenges.

4.2 Prioritizing practical application and real-world impact

Judges' feedback from the competition highlighted the critical role of practical application and real-world impact in projects, signaling a transformative shift in maker education towards applied learning. This focus is crucial for bridging the gap between academic theories and their practical applications in the real world, fostering a learning environment where students can see the direct impact of their innovations. As one judge aptly put it, "Projects that solve real-world problems not only demonstrate students' technical skills but also their understanding of societal needs." This insight is invaluable for educational institutions aiming to equip students with skills that extend beyond the classroom, ensuring that their learning experiences are directly relevant to real-world scenarios. By prioritizing projects with practical applications and societal

TABLE 1 The thematic analysis coding.

Themes (sections)	First-level code	Second-level code	Third-level code	Data source
4.1 Embracing transdisciplinary creativity	Integration of disciplines	Arts integration	Conceptual understanding	Feedback
			Transdisciplinary skills	Discussion
			Artistic expression	Interview
		Technology utilization	Creative technological solutions	Interview
			Technological proficiency	Feedback
			Innovative tech application	Discussion
4.2 Prioritizing practical application and real-world impact	Practical solutions	Addressing societal challenges	Real-world relevance	Interview
			Societal impact	Feedback
			Practical problem solving	Discussion
		Solution feasibility	Implementation viability	Interview
			User-oriented design	Discussion
			Solution sustainability	Discussion
4.3 Nurturing global awareness through sustainability	Environmental impact	Sustainable design principles	Environmental conservation	Discussion
			Sustainable practices	Interview
			Eco-friendly solutions	Feedback
		Global responsibility	Global impact	Discussion
			Ethical and responsible design	Interview
			Addressing global challenges	Discussion
4.4 Fostering cross-cultural collaboration and understanding	Cultural diversity	Teamwork across cultures	Diversity of perspectives	Feedback
			Effective communication	Interview
			Cultural exchange and learning	Interview
		Global relevance	Cultural sensitivity	Discussion
			Global innovation	Feedback
			Adapting to diverse viewpoints	Discussion
4.5 Encouraging adaptive problem-solving	Flexibility and resilience	Dynamic solution building	Responsiveness to challenges	Feedback
			Flexible thinking	Discussion
			Adapting to changing needs	Discussion
		Problem-solving strategies	Overcoming obstacles	Interview
			Innovative solutions	Feedback
			Creative problem-solving Approaches	Interview
4.6 Long-term impact on innovation mindset	Innovation mindset	Continuous learning	Lifelong learning	Feedback
			Curiosity and exploration	Discussion
			Embracing new knowledge	Interview
		Lifelong innovation	Risk-taking and experimentation	Discussion
			Adaptability in innovation	Discussion
			Persistent creative development	Interview

impacts, maker education can play a pivotal role in developing solutions to pressing global challenges, ultimately fostering a generation of students who are not just knowledgeable but also socially responsible and impact-driven.

- Promote real-world applications: emphasize the development of projects that address real-world challenges and societal needs.
- Bridge academic learning and practical impact: align maker education with practical applications, ensuring students' projects have tangible impacts.
- Cultivate solution-oriented mindsets: encourage an educational approach that nurtures solution-oriented thinking in students.

4.3 Nurturing global awareness through sustainability

The theme of sustainability in the judges' evaluations highlights the growing importance of global awareness and responsibility in maker education. Judges emphasized the need for projects to incorporate sustainable practices and consider their environmental impacts, aligning with the global movement towards more sustainable development. "Sustainable projects in maker education not only address environmental concerns but also teach students the importance of responsible innovation," one judge noted. This perspective is crucial in today's context, where environmental challenges require innovative solutions that are both effective and sustainable. By fostering a focus on sustainability in maker education, educators can prepare students to be conscientious global citizens who understand the importance of their impact on the world. This approach goes beyond traditional education, fostering a deeper sense of responsibility and ethical innovation in the next generation of makers.

- Incorporate sustainable practices: integrate sustainability into maker projects, teaching students to design with environmental consciousness.
- Teach global responsibility: educate students on the global impact of their projects, fostering a sense of ethical responsibility.
- Promote eco-friendly innovation: encourage the development of projects that are not only innovative but also beneficial to the environment.

4.4 Fostering cross-cultural collaboration and understanding

The judges' recognition of the value of cross-cultural collaboration in the competition underscores the necessity of preparing students for a globally interconnected world. Teams that harnessed diverse cultural perspectives were often able to produce more innovative and relevant solutions, illustrating the richness that diversity brings to problem-solving. A judge commented, "Diversity in teams brings a wealth of perspectives that often lead to more innovative outcomes." This insight stresses the importance of integrating cross-cultural collaboration in maker education, not only to enhance the creativity and scope of projects but also to foster understanding and respect among students from different cultural backgrounds. By promoting projects that encourage cultural exchange and global relevance, maker education can cultivate an environment that respects diversity, encourages inclusivity, and prepares students to operate effectively in a global context.

- Encourage cultural exchange: promote projects that bring together students from diverse cultural backgrounds, enhancing the richness of collaboration.
- Teach global relevance: ensure maker education includes a focus on developing globally relevant solutions.
- Nurture diverse perspectives: cultivate an educational environment that values and incorporates a variety of cultural viewpoints.

4.5 Encouraging adaptive problem-solving

The judges' appreciation for adaptive problem-solving skills in the competition highlights a critical skill set for the future of maker education. Their feedback emphasizes the importance of flexibility, resilience, and the ability to adapt solutions to new challenges or feedback. As one judge put it, "The ability to adapt and refine solutions is as important as the initial innovation." This perspective is particularly relevant in the rapidly changing modern world, where problems and technologies evolve quickly. By teaching students to be adaptable and responsive in their problem-solving approaches, maker education can foster a generation of innovators who are not only skilled but also agile and capable of navigating the complexities and uncertainties of the future.

- Develop flexible thinking: foster educational programs that emphasize adaptability and flexibility in problem-solving.
- Promote innovative solutions: encourage students to think innovatively and be open to evolving their projects.
- Teach resilience in design: incorporate resilience as a key component in maker education, preparing students to tackle unforeseen challenges.

4.6 Long-term impact on innovation mindset

The judges' insights reveal the significant role of maker competitions in cultivating a long-term mindset of innovation among participants. This aspect is crucial for sustaining a culture of creativity and exploration in maker education. A judge observed, "Fostering an enduring innovation mindset is crucial for the continuous evolution of ideas." This perspective highlights the need for maker education to go beyond temporary projects and foster a lasting focus on innovation. By encouraging continuous learning, exploration, and creative confidence, maker education can inspire students to pursue innovative endeavors throughout their lives, driving forward a culture of innovation and creative problem-solving.

- Foster continuous innovation: advocate for educational approaches that nurture a lasting focus on creativity and innovation.
- Encourage lifelong learning: promote opportunities for continuous learning and exploration beyond formal education.
- Inspire creative confidence: build programs that foster confidence in students to pursue creative and innovative endeavors.

5 Conclusions and future directions

The analysis of judges' feedback from the China-U.S. Young Maker Competition highlights key aspects for the evolution of maker education. Emphasis on transdisciplinary creativity, practical applications, sustainability, cross-cultural collaboration, adaptive problem-solving, and a long-term innovation mindset shapes the future of this educational approach. These insights offer a condensed roadmap for advancing maker education, highlighting the need for a comprehensive, experiential, and globally aware approach to equip students for future challenges. The findings point towards an educational shift to transdisciplinary creativity, integrating various disciplines to enhance problem-solving skills. The focus on practical applications with real-world impact reflects a move towards experiential learning that addresses societal challenges. Sustainability emerges as a crucial theme, aligning with global environmental consciousness. Cross-cultural collaboration is identified as key in preparing students for a globally interconnected world. Adaptive problem-solving is crucial for developing flexible and innovative thinkers. The study also highlights the importance of competitions in fostering a long-term innovation mindset, encouraging ongoing creativity and exploration.

Future directions include integrating maker education across various educational levels and disciplines, focusing on hands-on, real-world problem-solving. Incorporating entrepreneurship and business education within maker programs can bridge the gap between innovation and practical application. Promoting lifelong learning and innovation beyond formal education is essential for continuous skill development. Enhancing global connectivity and responsiveness to technological changes ensures that maker education remains relevant and forward-thinking. Continued research into the effectiveness and long-term impact of maker education will provide insights for its optimization.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Beijing Normal University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

WL: Conceptualization, Methodology, Writing – original draft. YZ: Formal analysis, Writing – review & editing. YiL: Conceptualization, Writing – original draft. ZF: Methodology, Writing – review & editing. YS: Writing – review & editing. XH: Formal analysis, Writing – original draft. YaL: Formal analysis, Writing – original draft. ML: Methodology, Supervision, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

We would like to express our heartfelt gratitude to every student and judge who have been part of the competition over the past decade. Their dedication, curiosity, and eagerness to learn have been a constant source of inspiration for us to enhance and refine our coaching approaches.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Al Mahmud, A., and Soysa, A. I. (2020). POMA: a tangible user interface to improve social and cognitive skills of Sri Lankan children with ASD. *Int. J. Hum. Comput. Stud.* 144:102486. doi: 10.1016/j.ijhcs.2020.102486

Asheim, B., Coenen, L., and Vang, J. (2007). Face-to-face, buzz, and knowledge bases: sociospatial implications for learning, innovation, and innovation policy. *Environ. Plan. C Govt. Policy* 25, 655–670. doi: 10.1068/c0648

Basadur, M., and Hausdorf, P. A. (1996). Measuring divergent thinking attitudes related to creative problem solving and innovation management. *Creat. Res. J.* 9, 21–32. doi: 10.1207/s15326934crj0901_3

Braun, V., and Clarke, V. (2023). Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. *Int. J. Transgen. Health* 24, 1–6. doi: 10.1080/26895269.2022.2129597

Carlsen, L., and Bruggemann, R. (2022). The 17 United Nations' sustainable development goals: a status by 2020. *Int. J. Sustain. Dev. World Ecol.* 29, 219–229. doi: 10.1080/13504509.2021.1948456

Chakraborty, S., Gonzalez-Triana, Y., Mendoza, J., and Galatro, D. (2023). Insights on mapping industry 4.0 and education 4.0. *Front. Educ.* 8:1150190. doi: 10.3389/feduc.2023.1150190

Chatterji, A., Glaeser, E., and Kerr, W. (2014). Clusters of entrepreneurship and innovation. *Innov. Policy Econ.* 14, 129–166. doi: 10.1086/674023

Clapp, E. P., Ross, J., Ryan, J. O., and Tishman, S. (2016). *Maker-centered learning: Empowering young people to shape their worlds* John Wiley & Sons.

de Bont, C. (2021). Furthering victor Papanek's legacy: a personal perspective. *J. Design Econ. Innov.* 7, 262–281. doi: 10.1016/j.sheji.2020.08.010

Desmet, P. M., Xue, H., Xin, X., and Liu, W. (2023). Demystifying emotion for designers: a five-day course based on seven fundamental principles. *Adv. Design Res.* 1, 50–62. doi: 10.1016/j.ijadr.2023.06.002

Dym, C. L., Agogino, A. M., Eris, O., Frey, D. D., and Leifer, L. J. (2005). Engineering design thinking, teaching, and learning. *J. Eng. Educ.* 94, 103–120. doi: 10.1002/j.2168-9830.2005.tb00832.x

Halverson, E. R., and Sheridan, K. (2014). The maker movement in education. *Harv. Educ. Rev.* 84, 495–504. doi: 10.17763/haer.84.4.34j1g68140382063

Hepp, A. (2020). The fragility of curating a pioneer community: deep mediatization and the spread of the quantified self and maker movements. *Int. J. Cult. Stud.* 23, 932–950. doi: 10.1177/1367877920922867

Kolb, D. A., Boyatzis, R. E., and Mainemelis, C. (2014). "Experiential learning theory: previous research and new directions" in *Perspectives on thinking, learning, and cognitive styles*. eds. R. J. Sternberg and L.-f. Zhang. (NJ: Lawrence Erlbaum), 227–247.

Lachheb, A., Abramenka-Lachheb, V., Moore, S., and Gray, C. (2023). The role of design ethics in maintaining students' privacy: a call to action to learning designers in higher education. *Br. J. Educ. Technol.* 54, 1653–1670. doi: 10.1111/bjet.13382

Lafont-Torio, J., Martín, J. M. M., Fernández, J. A. S., and Soriano, D. R. (2024). Perceptions of progress toward achieving the sustainable development goals: insights from cooperative managers. *Sustain. Technol. Entrep.* 3:100055. doi: 10.1016/j.stae.2023.100055

Lazonick, W. (2004). Indigenous innovation and economic development: lessons from China's leap into the information age. *Ind. Innov.* 11, 273–297. doi: 10.1080/1366271042000289360

Lindtner, S. (2015). Hacking with Chinese characteristics: the promises of the maker movement against China's manufacturing culture. *Sci. Technol. Hum. Values* 40, 854–879. doi: 10.1177/0162243915590861

Liu, W., Zhu, Y., Huang, R., Ohashi, T., Auernhammer, J., Zhang, X., et al. (2023). Designing interactive glazing through an engineering psychology approach: six augmented reality scenarios that envision future car human-machine interface. *Virtual Real. Intell. Hardw.* 5, 157–170. doi: 10.1016/j.vrih.2022.07.004

Liu, W., Zhu, Y., Liu, M., and Li, Y. (2021). Exploring maker innovation: a transdisciplinary engineering design perspective. *Sustain. For.* 14:295. doi: 10.3390/su14010295

Martin, L. (2015). The promise of the maker movement for education. *J. Pre Coll. Eng. Educ. Res.* 5:4. doi: 10.7771/2157-9288.1099

Martinez, S. L., and Stager, G. S. (2013). Invent to learn: makers in the classroom. *Educ. Digest* 79:11–15.

Miettinen, R. (2000). The concept of experiential learning and John Dewey's theory of reflective thought and action. *Int. J. Lifelong Educ.* 19, 54–72. doi: 10.1080/026013700293458

Ren, X., Silpasuwanchai, C., and Cahill, J. (2019). Human-engaged computing: the future of human–computer interaction. *CCF Trans. Pervasive Comput. Interact.* 1, 47–68. doi: 10.1007/s42486-019-00007-0

Sanders, E. B. N., and Stappers, P. J. (2012). *Convivial toolbox: generative research for the front end of design* BIS.

Tabarés, R., and Boni, A. (2023). Maker culture and its potential for STEM education. *Int. J. Technol. Des. Educ.* 33, 241–260. doi: 10.1007/s10798-021-09725-y

Tablatin, C. L. S., Casano, J. D., and Rodrigo, M. M. T. (2023). Using minecraft to cultivate student interest in STEM. *Front. Educ.* 8:1127984. doi: 10.3389/feduc.2023.1127984

Wen, H., Zhong, Q., and Lee, C. C. (2022). Digitalization, competition strategy and corporate innovation: evidence from Chinese manufacturing listed companies. *Int. Rev. Financ. Anal.* 82:102166. doi: 10.1016/j.irfa.2022.102166

Weng, X., Chiu, T. K., and Tsang, C. C. (2022). Promoting student creativity and entrepreneurship through real-world problem-based maker education. *Think. Skills Creat.* 45:101046. doi: 10.1016/j.tsc.2022.101046

Williamson, P. J. (2016). Building and leveraging dynamic capabilities: insights from accelerated innovation in China. *Glob. Strateg. J.* 6, 197–210. doi: 10.1002/gsj.1124

Yang, Z., Yu, C., Chen, X., Luo, J., and Shi, Y. (2022). Investigating user-defined flipping gestures for dual-display phones. *Int. J. Hum. Comput. Stud.* 163:102800. doi: 10.1016/j.ijhcs.2022.102800

Zhu, D., Al Mahmud, A., and Liu, W. (2024). Digital storytelling intervention for enhancing the social participation of people with mild cognitive impairment: co-design and usability study. *JMIR Aging* 7:e54138. doi: 10.2196/54138