Volume 27, Number 3 • Fall 2023 ISSN: 1087-3430

Electronic Journal for Research in Science & Mathematics Education

Flagship Journal of the International Consortium for Research in Science & Mathematics Education (ICRSME)

ICRSME

International Consortium for Research in
Science & Mathematics Education

The International Consortium for Research in Science & Mathematics Education (ICRSME) is seeking applications for Editor or Co-Editors for the *Electronic Journal for Research in Science & Mathematics Education* (*EJRSME*). The five-year term of editorship begins in 2024 and extends through 2028:

- 2024 - Incoming editors works alongside current editors
- 2025-2027 - Incoming editors assume full responsibility
- 2028 - Transition year for new editorial team

Transition years (first and fifth year) will ensure that incoming editors are familiar with all aspects of the online journal hosting system and are prepared for their role.

The ***Electronic Journal for Research in Science & Mathematics Education*** (***EJRSME***) is a peer-reviewed journal published quarterly by the International Consortium for Research in Science & Mathematics Education (ICRSME). EJRSME shares manuscripts relating to issues in science/mathematics education and science/mathematics teacher education from early childhood through the university level including informal science and environmental education. EJRSME reviews original science and mathematics education manuscripts that report meaningful research, present research methodology, develop theory, and explore new perspectives and teaching strategies.

To apply for the position, applicants must build an editorial team, consisting of, but not limited to, Editor(s), Managing Editor, and Copyeditor. The editorial team should have expertise in science/mathematics education or teacher education, as well as experience reviewing manuscripts for professional refereed educational journals. In addition, the editorial team will represent EJRSME at ICRSME Virtual Conferences and Consultations. The Editor(s) will meet quarterly with the Executive Directors of ICRSME.

Responsibilities of Editor(s):

- Maintain high academic standards for published manuscripts
- Publish four issues of the journal each year (Winter, Spring, Summer, Fall)
- Prepare and monitor reviewers for high quality
- Solicit manuscripts for the journal
- Screen submissions for relevance to journal and then assign manuscripts fitting the journal mission to an Associate Editor
- Consider reviews, decide disposition of manuscripts, and communicate decisions with authors
- Maintain communication between the journal and ICRSME
- Compile statistics and maintain files as appropriate for the journal
- Ensure that reviewers are thanked publicly on an annual basis
- Coordinate a "Publishing in *EJRSME*" session at ICRSME Consultations and Virtual Conferences

Full applications for the position are due by December 1, 2023 and must include:

- a cover letter of application that describes the applicant(s)' (1) experiences as an author, reviewer, or editor; (2) plan for building the editorial team (Managing Editor and Copy Editor); (3) commitment to and any prior involvement with ICRSME (consultations/virtual conferences, journal service, ICRSME newsletter, etc.); and (4) vision for the journal and operating procedures such as securing and preparing reviewers, and potential strategies for increasing visibility and possible revenue;
- full vita/vitae; and
- a statement detailing institutional support, if applicable.

Send application materials and questions to the current EJRSME Editors, Mark Bloom and Sarah Quebec Fuentes, at ICRSME.Consultation@gmail.com.

EJRSME
Electronic Journal for Research in Science & Mathematics Education

PUBLISHER
ICRSME, International Consortium for Research in Science & Mathematics Education

Cover designed by Dusty Crocker

CONTENTS

Volume 27 No. 3 | Fall 2023

RESEARCH / EMPIRICAL

ELECTRONIC JOURNAL FOR RESEARCH
IN SCIENCE & MATHEMATICS EDUCATION
VOL. 27, NO. 3, 1-19

Teachers' Evaluations of Geometry Problems That Use Visual Arts Contexts

Gloriana González
University of Illinois Urbana-Champaign

Christine N. Rinkenberger
Urbana School District #116

ABSTRACT

This study investigates geometry teachers' evaluations of problems that use visual arts contexts. We ask, *how do teachers' evaluations relate to the four arguments justifying the geometry course?* and *How do teachers draw on the APPRAISAL system to evaluate sample geometry problems from textbooks?* Nine high school teachers were convened in three focus groups. We analyzed the teachers' discussions using systemic functional linguistics and identified 676 evaluations. Ninety percent of the evaluations pertained to the system of appreciation, including teachers' reactions to the problems, their stances about the problems' compositions, and their opinions about the problems' values. The teachers valued problems where students could appreciate the relevance of math in real-world scenarios, engage in math explorations, and become intrigued.

Keywords: curriculum, geometry, art, real-world contexts, teachers' evaluations

Introduction

As a result of mathematics education reform efforts, curricular developers have aimed at designing math problems that will engage students in meaningful learning. Researchers in turn have analyzed the effects of problem-based curricula on students. For example, Boaler (1998) found that an open-ended approach enabled students to develop conceptual understanding and flexibility in solving a novel problem. Ridlon (2009) found that sixth-grade students who experienced a problem-based curriculum improved their achievement and their attitudes towards math. Cai et al. (2013) found that high school students who were exposed to a problem-based curriculum during middle school became good at problem posing. Nevertheless, the quest for research that yields a better understanding of how math curricula can support students' development of problem-solving skills continues to be relevant considering the educational challenges that the pandemic has made salient (Bakker et al., 2021). In the case of geometry instruction, there are curricular studies focusing on students' reasoning opportunities when solving textbook problems, especially proof tasks (Hunte, 2018; Otten et al., 2014). Some research studies also focus on teachers' implementation of the tasks (Sears & Chávez, 2014; Thompson & Senk, 2014). However, there are limited studies regarding teachers' perspectives about the geometry curriculum. Gooya's (2007) study about geometry teachers' perceptions regarding curricular reform in Iran is an exception. Developing an understanding of geometry teachers' perspectives about geometry problems in textbooks is critical for future work that attempts to change the content and the ways in which geometry is taught and learned in school.

In this article, we focus on investigating geometry teachers' perspectives about textbook geometry problems that are situated in visual arts contexts in the U.S. Math reform efforts call for curricular changes that would foster high school students' appreciation for the beauty of mathematics (National Council for Teachers of Mathematics [NCTM], 2018). Proponents of "STEAM" education have pushed to integrate science, technology, engineering, and math with art by extending the "STEM" acronym to include art (Stewart et al., 2021). STEAM-education initiatives emphasize problem-based and maker-space approaches (e.g., Lavicza et al., 2018; Martínez et al., 2022; Quigley & Herro, 2016; Quigley et al., 2020), which are well aligned with interdisciplinary approaches to math education (Bakker et al., 2021). Geometry instruction can offer special opportunities for students to develop such an appreciation for math in the world through connections with art, and U.S. geometry textbooks already attempt to show examples of problems situated in art-based contexts (González, 2020). These contexts, including drawing, architecture, and crafts, have the potential of increasing students' motivation and engagement. Moreover, the selection of problem contexts for students to reinvent math ideas, namely for *mathematizing*, is a key notion within the Realistic Mathematics Framework (Freudenthal, 1991).

Our goal in this study is to learn more about geometry teachers' perceptions and appreciations of geometry problems that integrate math and visual arts. While visual arts contexts have the potential for mathematizing, teachers' use of curricular materials can amplify or diminish this potential (Remillard, 2005). Therefore, understanding what teachers value about geometry problems situated in visual arts contexts can inform curriculum developers who intend to draw on teachers' perspectives in designing new curricula, and teacher educators who wish to build on teachers' knowledge to create learning opportunities to promote STEAM approaches. Theoretically, we rely on the "practical rationality of mathematics teaching" to unpack teachers' perspectives (Herbst & Chazan, 2003, p. 407). Methodologically, we apply linguistics and specifically appraisal theory, to analyze teachers' evaluations of geometry problems (Martin & White, 2005). We situate this study within the traditional justifications for the geometry course (González & Herbst, 2006) with an understanding that current discussions about new goals for the U.S. geometry curriculum do not happen in a vacuum and set expectations for why students should learn geometry.

Theoretical Considerations

The Practical Rationality of Mathematics Teaching

The notion of the practical rationality of mathematics teaching entails that teachers share an understanding of teaching practices related to specific content areas. For instance, geometry teachers in the U.S. have similar curricular resources and teach the same content. As a result, geometry teachers have shared perceptions and appreciations about teaching geometry. Drawing from Bourdieu's (1980) notion of "habitus," Herbst and Chazan (2003, p. 2) propose that teachers' "feel for the game" becomes explicit in conversations among teachers. Further elaboration of this construct has led to the identification of four professional obligations toward math teaching that guide teachers' decisions (Herbst & Balacheff, 2009; Herbst & Chazan, 2011). These obligations are toward the discipline of mathematics (*disciplinary*), *individual* student needs, developing and nurturing *interpersonal* relationships in the classroom, and school-specific rules and regulations (*institutional*). For example, geometry teachers emphasize the development of visualization skills related to geometric thinking (disciplinary obligation), while at the same time, they follow guidelines in their school math curriculum (institutional obligation). To elicit the practical rationality of mathematics teaching, researchers show videos or cartoon-based examples of instances of teaching that they have hypothesized to be typical (Herbst & Chazan, 2011). In teachers' discussions of these examples, they identify implicit norms in teaching that guide their decisions as well as what they value (or do not value) in teaching. The notion of the practical rationality of mathematics teaching facilitates identifying teachers' perspectives about instruction that make sense to mathematics teachers, although they may be difficult for others to understand, even mathematicians. According to Brousseau (1997), the didactical contract connects teachers, students, and content so that when implementing academic tasks in the curriculum, teachers can support student learning of specific content. The examination of the practical rationality of mathematics teaching centers on specific courses of study, such as an algebra or geometry course, so that the instructional demands when teaching specific content become explicit.

Geometry Instruction in the U.S.

Our focus on geometry teachers' perspectives about curricular materials is guided by historical considerations about the geometry curriculum that place special demands on teaching geometry. In the U.S., the geometry course has been a curricular requirement since the 1840s (Quast, 1968) and has overcome attempts to merge it with other courses (Stanic & Kilpatrick, 1992). A study of the justifications for the geometry course in the 20th century reveals that there are four arguments supporting its distinct contributions to the geometry curriculum (González & Herbst, 2006). The *mathematical argument* implies that geometry students will develop their reasoning skills as mathematicians. From this perspective, asking students to make mathematical conjectures and prove those conjectures is the main goal of the course. Proponents of the *intuitive argument* hold that geometry uniquely allows students to establish connections with real-world applications, and geometry curricula instills in students an appreciation for geometric patterns in their surroundings. The *formal argument* justifies the learning of geometry as valuable for developing logical reasoning that can be applied to everyday situations. The intent of teaching proofs using the two-column proof format is for students to apply logical reasoning. Finally, the *utilitarian argument* implies that the geometry course can provide knowledge and skills for students to apply to future employment. These arguments at times present conflicting values as they relate to curricular goals. For example, a geometry problem situated in the context of architecture could develop skills that architects apply in reading a floorplan, thus aligning the problem's context with the utilitarian argument. In contrast, by using geometry to study an architectural piece, students can come to appreciate geometrical patterns in the world, thus fulfilling the learning goals implied by the intuitive argument.

In our study, we focus on teachers' perceptions and appreciations of art-based geometry problems that are aligned with each of these four justifications for the geometry course. Geometry teachers' knowledge of the subject matter, the curriculum, and their students uniquely positions them to assess the potential value of using an art-based approach to geometry instruction. Since the teaching of geometry has been justified by competing discourses, teachers may hold various views about the purpose of learning geometry. Recent research pertaining to the U.S. Geometry curriculum has revealed the need to strengthen the connections between geometry and real-world applications (Desai et al., 2021). Considering new demands for changing the high school math curriculum, Geometry teachers' evaluations of textbook problems that are situated in an art-based context may elucidate their perspectives about these traditional justifications for the geometry course. By applying linguistic techniques for identifying appraisals, we elaborate how teachers' evaluations of art-based contexts used in geometry textbooks reveal teachers' perceptions and appreciations about teaching geometry when a curriculum includes competing justifications.

Teachers' Evaluative Stances and the APPRAISAL System

Systemic functional linguistics is a theory of language that proposes that speakers' meanings can be identified by the choices they make in their talk (Halliday & Matthiessen, 2004). To this end, systemic functional linguists aim to identify taxonomies that map speakers' choices according to the purpose of their talk. These choices are typically displayed as system networks that illustrate a system of linguistic choices. See Table 1 for this formation.

Table 1

System of APPRAISAL by Martin and White (2005)

Appreciations	Affects	Judgements
Reaction	Un/happiness	Normality
Composition	In/security	Capacity
Valuation	Dis/satisfaction	Tenacity
		Veracity
		Propriety

The theory is based on the principle that there are three overarching functions of language, called *metafunctions*, which are simultaneously accomplished in a text (oral or written). The *ideational metafunction* pertains to the goal of communicating experiences and focuses on the content of a text. The *interpersonal metafunction* is about establishing relations, including the function of proposing evaluations. The *textual metafunction* pertains to resources for organizing a text. Martin and White (2005) further develop an understanding of the APPRAISAL system, which describes speakers' choices when making evaluations in the English language. According to the theory, evaluations can be categorized according to the target of the evaluation, which constitutes the domain of ATTITUDE. Specifically, there are three subsystems: APPRECIATIONS for evaluating things, AFFECTS to show feelings, and JUDGEMENTS to evaluate people. Within each subsystem of ATTITUDES, there are other subsystems (see Table 1). The evaluations can be positive or negative, and speakers have more options in some domains to refine their evaluations than in others.

The theory of appraisal has been applied to various contexts, including evaluations in newspaper editorials (Achugar, 2004), communications in political discourse (Ross & Caldwell, 2020), tourist websites (Kaltenbacher, 2006), reflective prose in higher education assignments (Szenes, & Tilakaratna, 2021), and scientific reports (Stosic, 2021). Other studies have relied on systemic functional linguistics to analyze teachers' evaluations of teaching episodes by using other elements in the theory, such as modality (Herbst & Kosko, 2014; Kosko & Herbst, 2012). In this study, we apply Martin and White's (2005) approach to geometry teachers' discussions of problems from geometry textbooks. As artifacts of teaching and learning, textbook problems illustrate how the content of a discipline has been adapted to achieve the goals of schooling; a case of the *didactic transposition* (Chevallard, 1985). Our examination of geometry teachers' evaluations of geometry problems aims to identify the elements of the practical rationality of mathematics teaching that are at play when considering the possibility of integrating geometry and the visual arts. In alignment with the notion of the practical rationality of mathematics teaching (Herbst & Chazan, 2011), what teachers appraise demonstrates their categories of perception, since others without geometry teachers' specialized knowledge may be unable to discern the elements in a math problem that geometry teachers see. Teachers' evaluations constitute their categories of appreciation about what they value (or disregard) in geometry problems situated in visual arts contexts. Our investigation is guided by the goal of understanding what justifications geometry teachers value in problems situated in the visual arts so that future work that attempts to integrate math and the arts will consider teachers' perspectives.

Two research questions frame our examination of teachers' evaluative stances toward geometry problems situated in visual arts contexts. First, *how do teachers' evaluations relate to the four arguments justifying the geometry course (i.e., the formal, intuitive, mathematical, and utilitarian)?* Second, *how do teachers draw on the APPRAISAL system to evaluate geometry problems situated in visual arts contexts?* With the first question, we are interested in identifying the connection between what teachers perceive and what they value about geometry problems situated in visual arts in relation to the arguments justifying the geometry course. With the second question, we focus on the nature of the evaluations and whether and how teachers make use of various resources from the APPRAISAL system. Collectively, we are interested in learning more about aspects of the practical rationality of mathematics teaching, using the case of geometry teachers' evaluative stances toward a set of geometry textbook problems.

Methods

Participants and Data Collection

We conducted three focus group sessions with a total of nine high school geometry teachers. All teachers taught at public high schools in a state in the midwestern region of the U.S. that had adopted the Common Core State Standards for Mathematics and were recruited through announcements to their mathematics departments. The group had a wide range of teaching experience in the classroom, from novice (approximately three years) to veteran (over ten years). Each teacher had taught geometry during their career as a mathematics teacher. The focus group methodology allowed us to understand the practical rationality of mathematics teaching since teachers could share ideas that other teachers might deem acceptable (Herbst & Chazan, 2003). The two authors facilitated the sessions following the model described by Nachlieli (2011). Facilitator 1 (the second author) was in charge of managing the session, conducting the slide show, and asking the guiding questions. Facilitator 2 (the first author) asked probing questions to elicit and contrast various perspectives. Each session was video-recorded and transcribed for analysis. The participants were from seven different schools. See Table 2 for specific information about participants. School size varied from large to small schools.

Table 2

Focus Group Participants

Session	Participants	School	Approximate number of students
1	Curtis Maxwell	Violet HS	1,400
1	Gia Michaels	Honeydew HS	900
1	Charles Rankin	Umber HS	300
1	Emma Smith	Honeydew HS	900
1	Libby Walker	Violet HS	1,400
2	Chloe Baxter	Magenta HS	500
2	Renee Fedderly	Periwinkle HS	500
3	Skyler Beck	Catalina HS	1,400
3	Charity Oberlin	Byzantium HS	2,000

Note. Following institutional review guidelines, we use pseudonyms for participants and institutions.

Each session had four main parts. We started by framing the discussion around recent NCTM documents calling for revisions to the high school geometry curriculum and a clarification of terms pertaining to the strands of mathematical proficiency framework (Kilpatrick et al., 2001). Next, we engaged the participants in a 5- to 10-minute *wish list* activity, where we asked them to identify the characteristics of geometry lessons that promote mathematical proficiency through art and design. The participants had the opportunity to revise the list at the end of the session. In the third part of the session, we showed the teachers nine sample art-based problems from five different geometry textbooks aligned to the Common Core State Standard. See Table 3 for this information. The textbooks, published by mainstream publishers in the U.S., were selected based on the first author's study about the visual arts in geometry textbooks and represented various justifications for teaching geometry.

Table 3

Textbooks Used for Selecting Sample Problems

Acronym	Authors	Year	Title	Publisher
CME	Center for Mathematics Education Project	2013	*CME Geometry Common Core*	Pearson
CPM	Dietiker, L. & Kassarjian	2014	*Core connections Geometry*, 2nd *edition*	College Preparatory Mathematics
Glencoe	Carter et al.	2018	*Glencoe Geometry*	McGraw-Hill
Holt	Larson et al.	2012	*Geometry*	Holt McDougal
Pearson	Charles et al.	2015	*Geometry Common Core*	Pearson

The problems were typical short exercises where students are asked to apply their knowledge of geometric properties to a situation. In this case, we selected situations that involved art, such as a pottery design or an architectural structure. In the sessions, we presented the problems as they were written in the textbooks. We did not have examples of students' solutions to the problems or further information about the problems from the curricular materials. Our intention was for teachers to evaluate the problems with the same information that a student solving the textbook problems would have. The geometry teachers would apply their knowledge and experiences to their evaluations. Table 4 shows a description of the sample problems introduced.

Table 4

Descriptions of the Sample Problems

No.	Art Context	Math Topic	Standard	Argument	Textbook	Page No.
1	Sculpture	Right Triangles	HSG.SRT.B.5	Intuitive	Holt	449
2	Pottery	Circles	HSG.C.A.2	Utilitarian	Holt	707
3	Painting	Circles	HSG.C.B.5	Utilitarian	Glencoe	787
4	Calligraphy	Reflections	HSG.CO.A.5	Intuitive	Pearson	559
5	Drawing	Dilations	HSG.SRT.A.2	Intuitive	CME	323
6	Drawing	Congruence	HSG.SRT.B5 HSG.CO.C9	Mathematical Formal Intuitive	Glencoe	304
7	Architecture	Symmetry	HSG.CO.A.3	Mathematical Intuitive	Holt	616
8	Architecture	Pythagorean Theorem, Law of Cosines	HSG.SRT.B.5	Utilitarian	CPM	446
9	Architecture	Volume	HSG.GMD.A.3	Intuitive	Glencoe	813

Most of the problems were aligned with only one argument for justifying the geometry course, with the exceptions of problems six and seven. These problems had various parts that were aligned with different arguments; specifically, they required students to appreciate math in the world and comprised questions compelling students to complete a proof (formal argument) or propose a mathematical conjecture (mathematical argument). We selected these problems for the focus group sessions since problems with various underlying justifications are typical in geometry textbooks. In addition, there were limited problems in the textbooks aligned with the formal argument, and we wanted all arguments to be represented in the session. The participants answered the following guiding questions: (1) *Do you think that this problem is engaging for your students? Why or why not?* (2) *Does the context provide an entry point for students to learn math? How?* (3*) Is the problem promoting students' mathematical proficiency (i.e., conceptual understanding, procedural fluency, strategic competence, productive disposition, and adaptive reasoning)?* (4) *In what ways would this problem be valuable for teaching your students geometry?* (5) *Would you use this problem in your classroom? Why? If not, how would you adapt this problem?* After presenting each problem individually, we showed the problems in sets of three for the teachers to establish contrasts between them. In the last part of the session, the teachers analyzed the problem-based lessons that we created with different art-based contexts in mind. This analysis is beyond the scope of this article, and we report the results elsewhere (González et al., 2022).

Data Analysis

The sessions were fully transcribed. The transcripts show changes in turns of speech by the speakers, enumerated according to the sequence of participation in the talk. Following Martin and White (2005), the authors independently coded each turn by identifying (1) the appraising item (in bold), (2) the appraiser for that appraising item (italics), and (3) what was appraised (underlined). For example, 44 seconds after introducing problem four, Chloe said, "*I* think that <u>this one</u> is **kind of interesting** to think of someone who's always writing in a mirror image" (Session 2, Turn 117). Here, the appraising item is "kind of interesting," which is used to appraise "this one," a pronoun in reference to problem four. Chloe offered the evaluation, which is signaled when she said, "I think." The comment "to think of someone who's always writing in a mirror" provides a circumstance for considering this problem. Since what is appraised is a thing, the appraisal is an appreciation. Under appreciation, there are three subsystems: REACTION, COMPOSITION, and VALUATION. We coded this appraisal as "reaction," signaling that the problem grabbed her attention. In this case, "interesting" is a positive marker, although lessened by the modifier "kind of." The example illustrates Chloe's use of pronouns to refer to appraised items. In some cases, the appraiser or the appraised item was omitted or implied. For example, when discussing problem seven, Charity said, "So, **potentially engaging**" (Session 3, Turn 202). Here, it is unclear whether the appraiser was the teacher or if Charity implied that the problem could be engaging for the students, so we did not identify the appraiser. However, the appraised item could be recovered from the text since there was a discussion of problem seven.

In our analysis, we also considered *tokens* of evaluation. According to Martin and White (2005), appraisals could be inscribed or invoked. When an appraisal is inscribed, there is a clear link between the appraising item and what is appraised. In contrast, when an appraisal is invoked, there is an indirect connection. We found invoked appraisals when teachers used projected clauses to voice hypothetical classroom-based scenarios. For example, when discussing the three different parts of problem two, Charles said, "*I* really **hate** <u>A, B, C</u>. [Laughter.] Because **it** <u>takes all thought out of the process</u>. 'Here's the procedure.' 'What process to do you want me to memorize with the same problem on the test?' 'Here it is.' 'Memorize that.' 'I'm not going to think.'" (Turns 107-109). Here, "hate" describes a feeling, but it is a token of appreciation. That is, while the appraising item "hate" is a negative inscription of affect, the appraisal invokes a negative appreciation of the problem, a thing. The description of a hypothetical teacher-student exchange about asking for a procedure and giving a procedure in the quote is a token for an appreciation, showing a negative take on the composition of the problem because it lacks complexity. Following Martin and White (2005), we identified cases where projected speech was used to invoke evaluations as tokens because the speakers were invoking attitudes that, as analysts, we had to infer.

Another case where we identified tokens pertains to hybrids in the evaluations. According to Martin and White (2005), there is a hybrid when the inscribed and invoked attitudinal meanings differ. For example, in session one, during the teachers' discussion of problem one, Libby said that the problem was not engaging for students, and Curtis replied with the metaphor, "jump through this hoop." Libby then revoiced Curtis's comment and elaborated by stating, "Yeah. 'Jump through this hoop and show us now and independently master it.'" We identified two appraisals—"jump through this hoop" and "show us now" and "independently master it"—and we coded them as hybrids, combining JUDGEMENTS and APPRECIATIONS. The appraisals suggest a dual target for the evaluation. On the one hand, students need to show that they are capable of solving problems. Therefore, the evaluation is an inscription that makes a judgement about students' capabilities. On the other hand, the evaluation evokes a negative appreciation for the problem, specifically regarding its composition, since problem one lacks complexity. The comment suggests that according to the teachers, if students were to work on problem one, they would apply a procedure without necessarily showing their learning. The hybrid evaluations enabled the teachers in this example to offer an appraisal of problem one through the examination of the capabilities that students would need to solve the problem and demonstrate their learning.

We independently coded all of the transcripts and held subsequent meetings to resolve disagreements and refine the coding. We began by coding session one independently and checked the reliability of identifying appraising and appraised items. In addition, we checked if we agreed on the subsystem of evaluation (AFFECT, JUDGEMENT, or APPRECIATION), the subsystems within it, and if it was positive, negative, or a token. We agreed on 65% of the appraising items and 62% of the appraised items from the coding of the first session. When considering how many types of evaluations we had agreed upon, we found that we agreed on 87%. We realized that we had difficulties identifying appraising and appraised items. At times, it was difficult to recover textual references when speakers used pronouns to identify where tokens and hybrids took place. We continued to independently code the remaining sessions. Overall, when considering the appraising item, appraised item, and type of evaluation, the reliability was 72%, 69%, and 86%, respectively. We discussed our coding decisions, resolved disagreements, and reached a consensus.

Results

We start by reporting the findings regarding the evaluations that the teachers used in relation to the arguments justifying the geometry course to answer the first research question. Next, we answer the second research question by describing the resources from the subsystem of appreciation that the teachers used and their use of hybrid appraisals. Finally, we present the findings pertaining to the positive valuations that the teacher used to evaluate the problems, as these provide emerging evidence for what the teachers valued in the geometry problems situated in art-based contexts.

Teachers' Evaluations of Art-based Problems in Relation to the Arguments Justifying the Geometry Course

Overall, there seems to be evidence that the teachers preferred problems aligned with the utilitarian argument. Table 5 shows the results of all the appraisals offered by problem, aggregating the three sessions.[1]

[1] The results of appraisals per session are 217, 257, and 202 for sessions one, two, and three, respectively.

Table 5

Evaluations per Problem According to the System of APPRAISAL

Problem	Argument	Affect		Judgement		Appreciation		Total		
		Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg	Total
1	Intuitive	2	3	0	0	16	68	18 (20%)	**71 (80%)**	89
2	Utilitarian	3	2	2	1	75	26	**80 (73%)**	29 (27%)	109
3	Utilitarian	0	0	0	0	12	37	12 (24%)	**37 (76%)**	49
4	Intuitive	9	3	0	1	64	46	**73 (59%)**	50 (41%)	123
5	Intuitive	1	2	3	0	37	42	41 (48%)	**44 (52%)**	85
6	Mathematical, Formal, & Intuitive	0	2	0	1	19	44	19 (29%)	**47 (71%)**	66
7	Mathematical & Intuitive	0	0	0	0	13	14	13 (48%)	**14 (52%)**	27
8	Utilitarian	0	1	3	0	39	15	**42 (72%)**	16 (28%)	58
9	Intuitive	0	0	0	0	11	17	11 (39%)	**17 (61%)**	28
Other		0	1	0	1	9	4	9 (**60%**)	6 (40%)	15
Total		**15**	**14**	**8**	**4**	**295**	**313**	**318**	**331**	**649**

Note. "Other" refers to the general appraisals or combinations of appraisals toward more than one problem. "Pos" stands for positive, and "Neg" stands for negative. These results do not include hybrids.

The four problems with the highest number of evaluations were problem four (123 appraisals), problem two (109 appraisals), problem one (89 appraisals), and problem five (85 appraisals), which had the art-based contexts of calligraphy, pottery, sculpture, and drawing, respectively. Three of these four problems represented the intuitive argument (problems one, four, and five). Five problems (problems one, two, three, six, and eight) triggered evaluations with more than 70% positive or negative appraisals. Problems two and eight were evaluated mostly as positive (73% and 72% of the appraisals, respectively). These two problems represented the utilitarian argument. Specifically, problem two used the context of pottery for an archeologist to find the diameter of a plate by using a broken circular piece. Problem eight used the context of architecture, sharing the case of a person, Lashayia, who wishes to redesign a kitchen according to construction guidelines. In solving the problem, students would have to apply the Pythagorean theorem and trigonometry (i.e., the law of cosines) to determine whether the design meets the guidelines. In contrast, problems one, three, and six were mostly viewed as negative (80%, 76%, and 71% of the appraisals, respectively). These problems represented various arguments. Problem one concerned finding the height of a monument using trigonometry and represented the intuitive argument, as it was an example of how students could use geometry in their surroundings. Problem three was aligned with the utilitarian argument and shared the case of an artisan who had to rely on properties of circles to estimate the area of a mural. Problem six, situated in the context of drawing, would require students to examine a geometric figure that created an optical illusion. The problem represented the intuitive argument by asking students to appreciate the configuration of the visual arts piece. The problem uniquely represented the formal argument by asking students to complete a triangle congruence proof. Additionally, the problem requested students to explain their reasoning for establishing the relationship between two lines in the diagram, thus aligning the question with the mathematical argument. A further analysis of the evaluations proposed provides more nuance to the teachers' preferences.

Teachers' Uses of Resources from the Subsystem of APPRECIATION for Evaluating the Problems

The teachers used various resources from the system of appreciation to evaluate the problems, demonstrating complex analyses. Approximately one-third of the total appreciations offered by the teachers were from each subsystem of APPRECIATION, with slightly more appreciations coded as "composition." See Table 6 for this information.

Table 6

Evaluations of Problems According to the Subsystem of APPRECIATION

	Reaction			**Composition**			**Valuation**			Total
	Pos	**Neg**	**Total**	**Pos**	**Neg**	**Total**	**Pos**	**Neg**	**Total**	
Problem										
1	2	16	18	6	25	31	8	27	35	84
2	27	4	31	27	14	41	21	8	29	101
3	2	6	8	10	20	30	0	11	11	49
4	27	8	35	21	13	34	16	25	41	110
5	16	13	29	8	19	27	13	10	23	79
6	11	10	21	6	26	32	2	8	10	63
7	7	1	8	5	5	10	1	8	9	27
8	18	9	27	5	2	7	16	4	20	54
9	6	4	10	3	9	12	2	4	6	28
Other	2	1	3	1	0	1	6	3	9	13
Total	118 (**19%**)	72 (12%)	190 (31%)	92 (15%)	133 (**22%**)	225 (37%)	85 (14%)	108 (**18%**)	193 (32%)	608

Note. "Other" refers to the general appraisals or combinations of appraisals toward more than one problem. "Neg" stands for negative, and "Pos" stands for positive. These results do not include hybrids.

The findings show that the "reaction" appraisals were mostly positive. In contrast, the "composition" and "valuation" appraisals were mostly negative. These findings suggest that the teachers assumed a more critical stance through detailed analyses of the problems in terms of their characteristics and worth.

As an example of how the teachers' uses of resources from the system of appreciation allowed for a more sophisticated view of a problem, we discuss the teachers' evaluations of problem four. The problem, situated in the context of calligraphy, showed a Leonardo da Vinci illustration, where his handwriting appears in a mirror image. Students were asked to write the mirror image of the sentence, "Leonardo da Vinci was left-handed," and to discuss the possible reasons for his ease of writing mirror images of conventional text (Center for Mathematics Education Project, 2013, p. 559). There were a total of 110 appreciation appraisals for problem four. The reaction evaluations were mostly positive, with 27 positive reaction appraisals (25%) versus eight negative reaction appraisals (7%). The composition appraisals were also mostly positive, with 21 positive appraisals of composition (19%) and 13 negative appraisals of composition (12%). However, the valuation appraisals for problem four were mostly negative, with 25 negative appraisals of valuation (23%) and 16 positive appraisals of valuation (15%).

The positive reaction appraisals of problem four stated that the problem was "cool" "interesting, and "fun." The teachers also used tokens to state that the problem "is going to be intriguing" and that students "mostly heard of da Vinci." Some negative reactions were "why am I" writing backward? and that the problem was "never going to give you buy-in." With these negative evaluations, the teachers anticipated their students' reactions to the problem. The composition appraisals provided a more detailed evaluation of the characteristics of the problem. The positive composition appraisals included tokens to state that the problem "sneaks the math in" and "gets students thinking, writing." In terms of the problem's complexity, Libby stated that it was "something that anybody can try." Nevertheless, the composition appraisals were mostly negative. For example, the teachers stated that writing backwards would be "hard to do" and "a struggle." In addition, they stated that the problem "would take a lot of paper." The teachers noted that the statement of the problem did not include the required mathematical concepts to solve it. The teachers said, "not that it says reflection anywhere in that problem." Specifically, the problem did not discuss the concepts pertaining to reflections, such as the distance between the object and the mirror line, and instead stated "with it being that far from the line or that far." These examples show that by using composition appraisals, the teachers assessed specific characteristics of the problems, such as the language used or the mathematical concepts involved, thus revealing complex evaluations. Some positive valuations were that the problem could be "memorable," a "gateway," and "could lead you into talking about reflections." Nevertheless, some examples of negative valuation were "didn't really teach that much about symmetry," "dumbest thing ever," "not very mathematical," "is about being able to read it backwards," "more about being left handed," "the math's not really there," "wondering the math in it?," "not something that they're going to be graded on," and that writing backwards "might become a detriment later on." Skyler discussed how it was a problem "where you need a bunch of kids to do it and see what they do." The examples of evaluations toward problem four show how the teachers used various resources from the system of APPRECIATION to provide a multifaceted evaluation of the problems. Moreover, by using composition and valuation appraisals, the teachers changed their initially positive evaluation of a problem into a negative view with specific critiques.

Teachers' Uses of Hybrids for Evaluating the Problems

One characteristic of teachers' evaluations was the use of hybrids that combine APPRECIATIONS with AFFECTS or JUDGEMENTS. See Table 7 for this information.

Table 7

Hybrid Appraisals per Session

Session	Affect/Appreciation	Judgement/Appreciation	Total
1	8	4	12
2	11	3	14
3	1	0	1
Total	19	7	27

We found a total of 26 hybrid appraisals, mostly combining AFFECT and APPRECIATION. While the number of these types of appraisals is small, they speak to the nature of teachers' knowledge in terms of how they integrate their knowledge of their students into discussions about the mathematics curriculum. Session two was the session with the most hybrid appraisals (14), and session three was the session with the least hybrid appraisals (1). Hybrid appraisals of affect were more frequent than hybrid appraisals of judgement.

Teachers' hybrid appraisals of AFFECT and APPRECIATION mostly anticipated their students' feelings in relation to the problems. For example, when discussing problem six, Renee said that "*the kids* **hate** <u>triangles</u>," which we coded as dis/satisfaction and as a token of negative appreciation. In contrast, with regard to problem four, Charity said, "I think that *they* will be **curious**, that <u>it</u> actually works." We coded "curious" as a positive appraisal of satisfaction and as a token of positive appreciation for reaction. All of the hybrid appraisals of JUDGEMENT and APPRECIATION described capabilities, while also providing an appreciation for the problem. For example, when discussing writing backwards, Chloe said, "I **can't even do** that." We coded this appraisal as a token of negative judgement of capacity because of the apparent limitations in capabilities for writing backwards. At the same time, we coded this evaluation as a token of positive appreciation for composition, since there is suggestion that the process of writing backwards is complex. Earlier, we provided another example of a hybrid appraisal of judgement and appreciation with the phrase "jump through this hoop." Overall, the hybrid appraisals allowed the teachers to position themselves or their students in terms of their feelings or their characters when evaluating the problems.

Teachers' Positive Valuations of Problems Situated in Arts Contexts

By using reaction appreciations, the teachers stated their initial takes on the problems. By using composition appreciations, the teachers evaluated the sense of balance or the complexity of the problems. Ultimately, the valuation appreciations signaled whether, according to the teachers, the problems were worthwhile or not. That is, would the teachers keep or eliminate a problem and why? The teachers stated 85 positive valuation appraisals. We were interested in learning more about what was appraised with valuation. We listed the items that the teachers appraised, recovering the meanings from the transcription when they used pronouns. See Table 8 for these items and their appraisals.

Table 8

Positive Valuation Appraisals about Geometry Problems Situated in the Visual Arts

Item No.	Problem or alternative	Appraisal	Appraised
3	1A	might have been one that we went over	it [the problem similar to problem one]
204	1A	Like you don't have a choice	[doing word problems the whole day]
205	1A	you can't skip	it [doing word problems the whole day]
206	1	a little nice application of that	this [problem one]
207	1	put, on the worksheet about word problems	that [problem one]
264	1A	they're getting to do that themselves	[measuring the height of the statue in problem one]
268	1	matters	your context
269	1	matters	the measurements, when you're the one out there measuring the thing
42	2	really cool of discovery learning	this [problem two]
43	2	very different than, it's not just, "here's this figure, do the math skill that you do"	it [problem two]
63	2	you can't trace	crazy things that are really large
66	2A	I could incorporate that into the construction class	a problem like that
70	2A	integrity	the building
74	2	when you are talking about cross-curricular	[problem two]
75	2	introduced into like a world history class	having some of this [archaeology]
80	2	would like, more in a cross-curricular thing; where they could see the point better	it [archaeology]
210	2	like a real thing that could potentially happen	it [problem two]
212	2	an actual application that's like a real thing that a human would do normally	this [problem two]
214	2	using, to do the math	the pottery

Table 8

Positive Valuation Appraisals about Geometry Problems Situated in the Visual Arts

Item No.	Problem or alternative	Appraisal	Appraised
231	2	a good one to say, "When am I ever going to use this in life?"	it [problem 2]
232	2	"if you're an archaeologist, you might use it"	it [problem 2]
233	2	always a good utilist for some of those	it [problem 2]
235	2	maybe	that's what archaeologists do
240	2	cool	discussion [of different strategies for part "c"]
256	2	the math is situated in the actual use of the context	[problem 2]
438	2	see where that could be useful	that [problem 2]
439	2A	a little more of tying into something that they might want to do in the future	[a video of an archaeology dig]
440	2	might not have thought about, "Oh, I could use this for that"	[problem 2]
444	2	I could use	that [problem 2]
125	4	like	the concept of this problem
129	4	lets you open up to a lot of people that did stuff like this	it [da Vinci's reference]
185	4	gateway, to further mathematics	problem
186	4A	a compare and contrast	that [a problem with the word spelled backwards]
286	4	I see the mathematical part	in this [problem 4]
299	4	still would include	it [problem 4]
302	4	could lead you into talking about reflections	it [problem 4, part "a"]
306	4A	use, to talk about reflection	it [discussion of problem 4 at the beginning of the lesson]
307	4	does [provide an entry point for students to learn math], if it comes first, not #30	[problem 4]
311	4	memorable enough to remember	[problem 4]
312	4	could be a trigger.	[that da Vinci thing]
313	4	"Oh yeah!"	[problem 4]
465	4	could talk about reflection	[problem 4]
466	4	I think it does [promote mathematical proficiency]	[problem 4]
474	4	reflection is there	[problem 4]
545	4	I see, fitting	that [problem 4]
329	5A	"Ohhh, now I'm doing"	something with my art class and my Geometry class
335	5A	see	the connection [Geometry & art]
336	5A	might see, "Ohhh, they're related"	[Geometry & art]
486	5	get them [students] to think about what similar means	[problem 5]
487	5	could generate some good discussion	[what similar means]
488	5	there's a lot of exploration that could be done	two pictures [problem 5]
498	5	There's gotta be some math involved if you're looking at where the center of dilation is	[problem 5]
501	5	is definitely art design	it [problem 5]
502	5	is definitely used in logos, which is art	resizing an image
504	5	is drawing	it [problem 5]
535	5	there's more math going on in that one	number five

Table 8

Positive Valuation Appraisals about Geometry Problems Situated in the Visual Arts

Item No.	Problem or alternative	Appraisal	Appraised
538	5	has almost all of it	five
546	5	does for sure [have an opportunity to be a bigger problem]	five
517	6	is where you do the math	B [problem 6, part "b"]
518	6	is the naked math	there [problem 6, part "b"]
385	7A	an entry point	[a building known by students]
161	8	relevance	it [problem 8]
162	8	relevant	it [problem 8]
165	8	"Oh, I can see value in why you want to solve"	it [a problem with buy-in]
166	8	more relevance	[problem 8]
387	8	a real-world application	[problem 8]
390	8	interested	construction and building houses and stuff
391	8	"Oh yeah, I know that about this. And now I'm going to solve this triangle to see if it really does fit"	[problem 8]
392	8	I know	the math behind it [problem 8]
394	8	that you could use later to engage the kids	a question [problem 8]
398	8	more prevalent	blueprints
400	8	be good	this [problem 8]
573	8	at least shows that math is used somewhere	it [problem 8]
574	8	a good thing	that [showing that math is used somewhere]
581	8	a little less forced though than some of the other like, quote—unquote, real-world applications	it [problem 8]
584	8	less forced	This [problem 8 vs. another problem about making a ramp following building code]
585	8	I don't think is forced	this one [problem 8]
415	9	more cultural	problem [9]
596	9	want	kids to find the volume of a cone
260	1 & 3	you're actually doing math, I guess	[problems 1 & 3]
261	1 & 3	you're like using numbers and stuff	[problems 1 & 3]
361	4 & 6	have more of an entry point than five	[problems 4 & 6]
362	4 & 6	useful	the context [problems 4 & 6]
583	general	No [does not sound forced]	[another problem about making a ramp following building code]
587	general	"hey there's a ramp"	[another problem about making a ramp following building code]

Note. The appreciation appraisals are numbered by session as they appear in the transcripts: appraisals 1–187 pertain to session one, appraisals 188–418 pertain to session two, and appraisals 419–608 pertain to session three. We use "A" to denote when the appraised item is an alternative to the problem versus the provided item. Appraisals pertaining to usefulness are bolded, to mathematics are italicized, and to discovery are underlined.

The list of appraised items using positive valuation appraisals includes 13 items that are alternative to the problems provided (15% of the positive valuation appraised items). This is relevant because at times the teachers altered the problems that we provided them to discuss the characteristics that they would value in problems. Thirty of the positive valuation appraisals (35% of the positive valuation appraised items) pertained to statements regarding the opportunity to use math in real-world situations. These appraisals were aligned with the utilitarian argument. Eighteen of the positive valuation appraisals (21% of the positive valuation appraised items) pertained to the mathematical content of a problem. For example, the teachers named specific valuable content (e.g., the center of dilation in appraisal 498) or how a problem provides an entry point to the mathematical ideas in a lesson (e.g., reflection in appraisal 402). The attention to the mathematical content of the problem was aligned with the mathematical argument. Six appraisals (7% of the positive valuation appraised items) revealed that the teachers valued problems where students can discover a new idea and apply multiple solution strategies. The attention to discovery-oriented opportunities was aligned with the intuitive argument. Altogether, the teachers' evaluations were related to pedagogical decisions regarding how students come to learn a new idea through problem solving.

The teachers' positive valuation appraisals also included comments regarding the desirable characteristics of geometry problems that are situated in visual arts contexts. Namely, the teachers valued problems that coherently connected the mathematical content and the problem's context (e.g., "a little less forced though than some of the other like, quote—unquote, real-world applications" appraisal 581). The teachers also valued opportunities to use a context to motivate students (e.g., "could be a trigger," appraisal 312). The teachers saw the promise in adapting problems where students would use their knowledge of their surroundings (e.g., choosing a building known to the students as an example of architecture, appraisal 385) and interests (e.g., "construction and building houses and stuff" are contexts for students' interests, appraisal 390). The teachers also stated that cross-curricular problems (e.g., "would like it more in a cross-curricular thing; where they could see the point better," appraisal 80) and problems that include cultural connections (e.g., "more cultural," appraisal 415) are valuable. Overall, the positive valuation appraisals showed that in evaluating the problems, teachers contended with a problem's contexts, the mathematical ideas in the problem, and the pedagogical aspects of how a problem provides opportunities for students to learn geometry. The teachers' valuations revealed elements of the practical rationality of mathematics teaching by illustrating what teachers perceive and appreciate.

Discussion

According to the teachers in this study, the integration of geometry and art is possible. Nevertheless, their appraisals of the sample problems yielded a complex picture of what they value in geometry problems. These appraisals revealed the practical rationality of mathematics teaching in that the teachers considered how the problems target specific math concepts, thus showing their responsibility to portray the discipline of mathematics. Additionally, their appraisals revealed their anticipations of their students' feelings towards the problems—students' likes, dislikes, motivation, curiosity—as well as their capabilities. In the teachers' consideration of their students' perspectives, they were revealing their obligation towards their individual students, which is a component of the practical rationality of mathematics teaching. Additionally, there is an interpersonal component in the teachers' attention to students' feelings because they considered students' motivation and engagement during problem-solving. The teachers' suggestions to adapt problems to their students' local context is another example of how they strived to attend to their individual students' interests. While there was less evidence of how the teachers attended to the institutional obligation when reviewing and making curricular choices, all of the problems were aligned with the Standards as established in their school curriculum. The various "obligations" that mathematics teachers have to fulfill include teaching mathematics and taking care of their students (Bieda et al., 2015; Herbst & Balacheff, 2009; Herbst & Chazan, 2011). Our study is consistent, showing how in geometry teachers' evaluations of problems situated in art-based contexts, they attended to the mathematical content to fulfill their disciplinary obligations and to their students' interests and needs to fulfill their obligations toward individual students. We were able to elicit the practical rationality of teaching by holding discussions of curricular materials and learning more from teachers about what they find appropriate to use in their classroom and why.

In terms of the arguments justifying the geometry, the teachers evaluated the sample problems aligned with the utilitarian argument as valuable. This finding is relevant considering current work that seeks to connect math and design in STEAM education (Bush et al., 2018). It seems that the problems that are aligned with the utilitarian argument would help teachers to answer a question that students often ask, "Why do I need to learn this?" The sample problems with art-based contexts that were embedded in jobs such as being an archeologist, a painter, or an architect positioned the students as someone who must use geometry in their professional practice. Curricular designers who are seeking authentic opportunities for students to appreciate geometry may want to investigate contexts aligned with the utilitarian argument. At the same time, the teachers had positive evaluations towards problems aligned with the intuitive argument for various reasons, including the opportunity to engage in discovery-oriented investigations as well as the chance to appreciate math in the world. Additionally, the teachers' attention to the math content of the problems (or their lack of math content), whether an art-based context provided students an entry point to examine math ideas, and the opportunity for students to display various solution strategies to a math problem were aligned with the mathematical argument. In contrast, there was limited evidence of teachers' preference for problems aligned with the formal argument. In their discussions of the problems, the teachers did not seem to argue in favor of problems that promote opportunities for *doing proofs*, which typically rely on a two-column format for writing statements and reasons that justify the solution of a problem (Herbst, 2002), as one of the goals of the geometry course. Therefore, it seems that teachers are less fond of proof tasks that appear in the geometry curriculum (Otten et al., 2014). Further investigation is needed to see whether and how the formal argument justifying the geometry course is one that teachers continue to support.

Overall, the teachers demonstrated a sophisticated analysis of the sample geometry problems. The linguistic methods revealed a complex picture of their evaluations. With REACTION, the teachers anticipated what their students would say about the problems. With COMPOSITION, the teachers analyzed the problems' complexity and coherence. Then with VALUATION, the teachers showed worthwhile characteristics of the problems. The teachers' evaluations were sophisticated and unpacked a deep analysis of the problems. For example, the teachers valued coherence between the math content and the art-based context. In contrast, the teachers critiqued cultural contexts that trivialized the authenticity of the problem. Therefore, teachers' involvement in designing geometry problems that use art-based contexts would bring attention to important issues that are close to their students.

Conclusion

Many voices are calling for changes to the geometry curriculum (NCTM, 2018) and for investigating interdisciplinary approaches to math instruction (Bakker et al., 2021). There are suggestions for a new geometry curriculum that can leverage students' experiences and use geometry to represent their noticings and wonderings (Desai et al., 2021). The proponents of STEAM have looked into connections between math and art, and most recently, using design thinking as a unifying theme (Bush et al., 2018). With a design thinking theme, curricular developers may be able to target the utilitarian argument by extending students' use of geometry concepts to authentic problems. At the same time, students may be drawing on their intuition to see geometry embedded in real-life spaces and situations.

Geometry teachers possess knowledge of math and their students that become crucial in curricular adaptations. According to the Standards by the Association of Mathematics Teacher Educators (AMTE, 2017) in the U.S., well prepared teachers can anticipate students' thinking, and are also knowledgeable of contexts that shape students' learning. Professional development initiatives can build on that knowledge and promote connections between geometry and art. For example, Verner et al. (2019) report on an initiative in Israel for teachers to use ethnomathematics in crafting tasks for their students to learn more about geometry with cultural artefacts. Work on ethnomodeling also shows the potential for incorporating culturally sustaining teaching practices in geometry (Desai et al., 2022).

Methodologically, our study provides insights about teachers' linguistic choices when making evaluations about curricular materials with their peers. It may be relevant to investigate prospective teachers' evaluations, mirroring other work that focuses on their notions of *good* problems (e.g., Crespo & Sinclair, 2008). We recognize the various limitations of our study, including the small number of participants and the limited number of problems that we showed to the teachers. Nevertheless, our study does not concern a particular curricular approach but rather the more general use of the visual arts in relation to the traditional justifications for the geometry course. To make sustainable changes to the geometry curriculum, it is crucial to understand teachers' perspectives. The quest for meaningful contexts for students to learn and enjoy mathematics can benefit from teachers' knowledge of their students, the curriculum, and mathematics.

This work was supported by the by the Office of the Vice Chancellor for Research at the University of Illinois Urbana-Champaign through a Campus Research Board award to Gloriana González for the project GeoDesiC-Geometry and Design in Classrooms [RB19058].

Gloriana González (ggonzlz@illinois.edu) is a Professor in the Curriculum & Instruction Department at the University of Illinois Urbana-Champaign and University of Illinois Scholar. Dr. González is currently investigating lesson study adaptations to support teachers' implementation of a human-centered design approach for geometry problem-based instruction. She was a middle school and high school teacher in Massachusetts and Puerto Rico.

Christine N. Rinkenberger (crinkenberger@usd116.org) is an assistant principal of Curriculum-Instruction at Urbana High School. She has a master's degree in Curriculum Instruction, as well as Educational Policy from the University of Illinois Urbana-Champaign. Christine was a high school math and ESL teacher prior to becoming a school administrator.

References

Achugar, M. (2004). The events and actors of 11 September 2001 as seen from Uruguay: Analysis of daily newspaper editorials. *Discourse and Society, 15*(2-3), 291-320. https://doi.org/10.1177/0957926504041021

Bakker, A., Cai, J., & Zenger, L. (2021). Future themes of mathematics education research: An international survey before and during the pandemic. *Educational Studies in Mathematics*, *107*(1), 1-24. https://doi.org/10.1007/s10649-021-10049-w

Bieda, K., Sela, H., & Chazan, D. (2015). "You are learning well my dear": Shifts in novice teachers' talk about teaching during their internship. *Journal of Teacher Education, 66*(2), 150-169. https://doi.org/10.1177/0022487114560645

Boaler, J. (2008). *What's math got to do with it?* Penguin Books.

Brousseau, G. (1997). *Theory of didactical situations in mathematics: Didactique des mathématiques 1970-1990* (N. Balacheff, M. Cooper, R. Sutherland, and V. Warfield, Eds. and Trans.). Kluwer.

Bush, S. B., Cox, R., & Cook, K. L. (2018). A critical focus on the M in STEAM. *Teaching Children Mathematics*, *23*(2), 110-114.

Carter, J. A., Cuevas, G. J., Day, R., & Mallory, C. (2018). *Glencoe geometry.* McGraw-Hill.

Center for Mathematics Education Project. (2013). *Geometry.* Pearson.

Charles, R. I., Hall, B., Kennedy, D., Bass, L., Johnson, A., Murphy, S. J., & Wiggins, G. (2015). *Geometry common core.* Pearson.

Chevallard, Y. (1985). *La transposition didactique: Du savoir savant au savoir enseigné*. La Pensée Sauvage.

Crespo, S., & Sinclair, N. (2008). What makes a problem mathematically interesting? Inviting prospective teachers to pose better problems. *Journal of Mathematics Teacher Education 11*(5), 395-415. https://doi.org/10.1007/s10857-008-9081-0

Desai, S., Bush, S.B., & Safi, F. (2021). Mathematical representations in the teaching and learning of geometry: A review of the literature from the United States. *Electronic Journal for Research in Science & Mathematics Education, 25*(4), 6-22. https://ejrsme.icrsme.com/article/view/20634

Dietiker, L. & Kassarjian, M. (2014). *Core connections geometry* (Second Edition, Version 5.0). CPM Educational Program.

Freudenthal, H. (1991). *Revisiting mathematics education: China lectures.* Kluwer.

González, G. (2020). Visual art contexts in U.S. geometry textbooks aligned with the Common Core Standards. *Annales de Didactique et de Sciences Cognitives, 25,* 183-210. https://doi.org/10.4000/adsc.560

González, G., & Herbst, P. (2006). Competing arguments for the geometry course: Why were American high school students to study geometry in the twentieth century? *International Journal for the History of Mathematics Education, 1*(1), 7-33.

González, G., Kim, G.-Y., & Rinkenberger, C. (2022, April). Geometry teachers' perspectives on problem-based lessons situated in arts-based contexts. Paper accepted to be presented at the American Educational Research Association annual meeting, Chicago, IL

Halliday, M. A., K., & Matthiessen, C. (2004). *An Introduction to Functional Grammar* (3rd ed.). Hodder Arnold.

Herbst, P. G. (2002). Engaging students in proving: A double bind on the teacher. *Journal for Research in Mathematics Education*, *33*(3), 176-203. https://doi.org/10.2307/749724

Herbst, P., & Balacheff, N. (2009). Proving and knowing in public. In M. Blanton, D. Stylianou, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. 40-63). New York, NY: Routledge.

Herbst, P., & Chazan, D. (2003). Exploring the practical rationality of mathematics teaching through conversations about videotaped episodes: the case of engaging students in proving. *For the Learning of Mathematics, 23*(1), 2-14. https://www.jstor.org/stable/40248404

Herbst, P., & Chazan, D. (2011). Research on practical rationality: Studying the justification of action in mathematics teaching. *The Mathematics Enthusiast*, *8*(3), 405-462. https://doi.org/10.54870/1551-3440.1225

Herbst, P., & Kosko, K. W. (2014). Using representations of practice to elicit mathematics teachers' tacit knowledge of practice: a comparison of responses to animations and videos. *Journal of Mathematics Teacher Education*, *17*(6), 515-537. https://doi.org/10.1007/s10857-013-9267-y

Kaltenbacher, M. (2006). Culture related linguistic differences in tourist websites: The emotive and the factual. In G. Thompson & S. Hunston (Eds.), *System and corpus: Exploring connections* (pp. 269-292). Equinnox.

Kilpatrick, J., Swafford J., & Findell B. (Eds.). (2001). *Adding it up.* National Academy Press.

Kosko, K. W., & Herbst, P. (2012). A deeper look at how teachers say what they say: A quantitative modality analysis of teacher-to-teacher talk. *Teaching and Teacher Education*, *28*(4), 589-598. https://doi.org/10.1016/j.tate.2011.11.010

Larson, R., Boswell, L., Kanold, T. D., & Stiff, L. (2012). *Geometry.* Holt McDougal/Houghton Mifflin Harcourt.

Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English.* Palgrave Macmillan.

Nachlieli, T. (2011). Co-facilitation of study groups around animated scenes: The discourse of a moderator and a researcher. *ZDM*, *43*(1), 53-64. https://doi.org/10.1007/s11858-010-0305-2

National Council of Teachers of Mathematics (NCTM). (2018). *Catalyzing change in high school mathematics*. NCTM.

Quast, W. G. (1968). *Geometry in the high schools of the United States: An historical analysis from 1890 to 1966*. Unpublished doctoral dissertation. Rutgers—The State University of New Jersey, New Brunswick.

Ross, A. S., & Caldwell, D. (2020). 'Going negative': An appraisal analysis of the rhetoric of Donald Trump on Twitter. *Language & communication*, *70*, 13-27. https://doi.org/10.1016/j.langcom.2019.09.003

Stanic, G. M., & Kilpatrick, J. (1992). Mathematics curriculum reform in the United States: A historical perspective. *International Journal of Educational Research*, *17*(5), 407-417. https://doi.org/10.1016/S0883-0355(05)80002-3

Stewart, A. J., Mueller, M., & Tippens, D. J. (2021). *Converting STEM into STEAM programs*. Springer.

Stosic, D. (2021). Persuasion strategies for demonstrating topic significance in reports of randomised controlled trials. *English for Specific Purposes, 62,* 1-14. https://doi.org/10.1016/j.esp.2020.11.002

Szenes, E., & Tilakaratna, N. (2021). Deconstructing critical reflections in social work and business: Negotiating emotions and opinions in reflective writing. *Journal of English for Academic Purposes, 49*. https://doi.org/10.1016/j.jeap.2020.100931

ELECTRONIC JOURNAL FOR RESEARCH
IN SCIENCE & MATHEMATICS EDUCATION
VOL. 27, NO. 3, 20-34

Utilizing Cognitive Load Theory and Bruner's Levels of Developmental Learning to Address Students' Struggles Related to Area of Polygons: A Pedagogical Action Research Study

Beth Cory
Sam Houston State University

Amy Ray
Sam Houston State University

ABSTRACT

In this pedagogical action research study, we, as post-secondary mathematics teacher educators, built on an existing effort to improve pre-service teachers' mathematical vocabulary understandings by intentionally addressing their struggles related to polygonal area formulas. Utilizing cognitive load theory and Bruner's levels of developmental learning, we adapted and refined an existing "Area of Polygons" lesson to eliminate extraneous elements and scaffold the introduction of essential elements in the context of a cognitively engaging activity. Comparing our resulting lesson components to existing literature on polygonal area, we found two main approaches towards exploring area of polygons. Both approaches emphasized conservation of polygonal area with one focused on the details of attributes and square units and the other focused on comparisons of areas of figures. We discuss the implications of these approaches and the use of cognitive load theory in tandem with Bruner's levels for future curriculum redesign.

Keywords: geometry, pedagogical action research, pre-service teacher education

Introduction

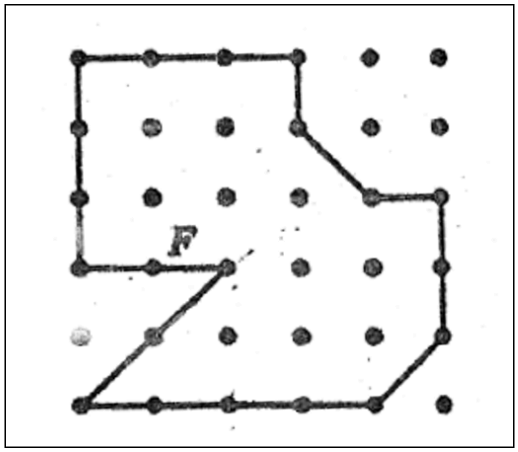
We, as post-secondary mathematics teacher educators, often collaboratively consider the content that our students, elementary pre-service teachers (PSTs), find difficult across the sequence of mathematics courses we teach. As part of a larger study focused on instructor collaboration and pedagogical improvements across the course sequence, we found that our PSTs routinely struggle with mathematical vocabulary (Bullock et al., 2021). Building off this work and connecting to concepts of cognitive load theory (Ayres, 2006; Pass et al., 2003) and Bruner's levels of developmental learning (Bruner, 1986; Reys et al., 2012), the current pedagogical action research study explores our response to students' struggle with polygonal area formulas. The impetus for this study, noticings about students' misconceptions about polygonal area and area formulas, is described below.

In Math 1385: Foundations of Mathematics II, the second mathematics content course in the sequence focused on geometry and measurement, one important topic that students explore is polygonal area formulas. In the original version of the "Area of Polygons" lesson, the first author spent about an hour carefully guiding students through interactive activities with patty paper to help them understand the derivations of the rectangle, parallelogram, triangle, and trapezoid area formulas.

At the end of the lesson, she had the students apply their new knowledge by finding the areas of "Crazy Shapes," an activity adapted from Aichele and Wolf (2008, p. 18). An example figure is provided in Figure 1.

Figure 1

Crazy Shape (Aichele & Wolf, 2008, p. 18).



She anticipated that students would divide the large crazy shape into smaller squares, rectangles, parallelograms, triangle, or trapezoids, find their areas, then add these areas together to find the area of the large crazy shape. However, as she walked around the room, she noticed one student was partially dividing the crazy shape into triangles, using the formula $b \times h$ to find the triangles' areas. Interested in this student's thought process, the first author initiated the following discussion:

> Cory: Why $b \times h$ for the area of a triangle?
> Student: Because that's how we find the area of triangle!
> Cory: How does the area of a triangle relate to the area of a parallelogram?
> [Extended silence.]
> Cory: Do you remember that a triangle's area is always half a parallelogram's area?
> [Extended silence.]
> Cory: So, what is the area formula for a triangle?
> Student: It's always base times height.

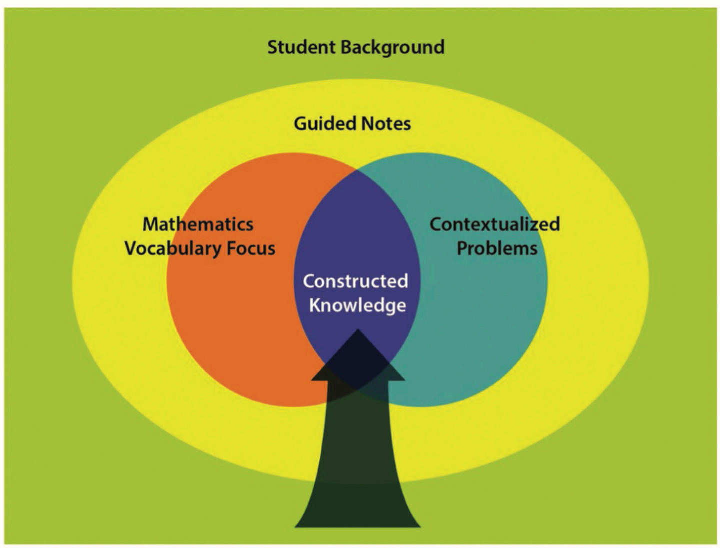
After teaching the "Area of Polygons" lesson a few more times in other sections of the course, the first author noticed similar responses from a number of students. These students seemed to be associating area with "base times height" or "length times width" no matter the type of shape. The first author's experiences initiated a series of lesson revisions intended to better enable students to grasp and retain an understanding of the connections between polygonal area formulas, specifically targeting the $\frac{1}{2}$ in the triangle area formula.

Study Background and Guiding Conceptual Framework

In our work with elementary PSTs, our research team, composed of five mathematics teacher educators (MTEs), embarked on a multi-year effort to improve learning opportunities for our students using guided notes across the sequence of mathematics courses for future elementary teachers. At our Southeastern, public university with a Hispanic Serving Institution (HSI) designation, our student population includes 45% first-generation undergraduates and over 75% of the students are employed while pursuing their degree. Our PSTs take three required elementary foundations mathematics content courses in the mathematics department and one mathematics methods course in the education department. Our team, initially consisting of three mathematics faculty and two education faculty, began our curriculum improvement efforts in response to observations about students' struggles that we routinely observed semester to semester across these courses. These efforts are summarized in the Guided Mathematics Vocabulary (GMaV) Conceptual Framework in which we position students' constructed knowledge at the intersection of an explicit focus on mathematics vocabulary and the use of contextualized problems using the curriculum tool of guided notes. See Figure 2 for this information.

Figure 2

GMaV Conceptual Framework (Bullock et al., 2021).



We acknowledge, seek to understand, and account for the mathematical knowledge that our students bring to our classrooms as we strive to support students' construction of knowledge. Specific insight into these ongoing curriculum and assessment reform efforts can be read in more detail elsewhere (Bullock et al., 2021; Ray et al., 2023).

As a part of our curriculum and assessment reform work, we focused on key topics or lessons that could be refined and adjusted to better support students' mathematical vocabulary understandings and, more broadly, address common areas of misconceptions or difficulty. One specific lesson that stood out to us in MATH 1385 was the "Area of Polygons" lesson (hereafter referred to as the "lesson"). We wanted our students to "look for and make use of structure" (CCSSI, 2010, p. 8) as they make connections between the polygonal area formulas. This Standard for Mathematics Practice (SMP) detailed in the Common Core State Standards for Mathematics (CCSSM), emphasizes the importance of recognizing and analyzing patterns and structures of mathematical objects. Within the context of polygonal area, we often noticed students arriving to MATH 1385 with a memory of area involving the multiplication of attributes but lacking a robust conceptual understanding of polygonal area. Additionally, even after the original lesson, students were not leveraging connections between the polygonal area formulas to complete higher level tasks.

Relevant Research Literature

Cognitive Load Theory

As we worked to develop and refine our lesson, we utilized various aspects of cognitive load theory. According to cognitive load researchers (Ayres, 2006; Pass et al., 2003), learners have a limited working memory, the space in the brain where all conscious cognitive processing occurs. Researchers posit that working memory may only be able to handle two to three new ideas, or elements, at a time. Thus, the instructor's goal is to load as little of the students' working memory space as possible so that novel ideas can be learned efficiently. As a student learns the knowledge elements associated with a particular mathematical concept, the student may incorporate (and perhaps automate) those elements into what is called a *schema*, or web of ideas. Once a schema is formed in a student's long-term memory, their working memory can more easily process related novel ideas and hook them into the already existing schema. Ultimately, a student's schema is so well-learned that it begins to act like a single element, thus vastly expanding the processing capability of working memory.

As a student learns a new mathematics concept, three types of cognitive load may use up the space in working memory: a) intrinsic, b) germane, and c) extraneous. *Intrinsic load* results from the interactivity of elements that are essential or intrinsic to understanding a certain mathematical concept. For example, to understand the triangle area formula, a student might need to understand the elemental definitions of area, triangle, parallelogram, base, height, the multiplication operation, as well as the fractional concept of $\frac{1}{2}$. These elements interact together to create the triangle area formula. The more elements necessary for understanding a concept, the higher the intrinsic load will be. *Germane load* refers to the cognitive activity necessary for a student to form a schema from the necessary knowledge elements. An instructor might influence germane load by creating, for example, activities that involve productive struggle, thus providing students opportunities to engage deeply with the new elements involved in the mathematics concept. *Extraneous load* involves unnecessary cognitive activity resulting from the way the instructor or textbook presents the information. For example, instructional materials may misdirect attention to nonessential aspects of the concept or needlessly require learners to search for relevant information.

According to cognitive load theory, the instructor's goal is to eliminate extraneous load and to decrease intrinsic load through carefully timing and scaffolding the introduction of essential elements. This in turn frees up a student's working memory for a higher germane load so that a robust schema can be developed.

Bruner's Levels of Developmental Learning

We posit that one way to scaffold the introduction of intrinsic knowledge elements so as not to overload working memory, and also guide the instructor in creating activities that boost an appropriate level of germane load, is to incorporate Bruner's levels of developmental learning. These three progressive levels are 1) enactive, 2) iconic, and 3) symbolic (Bruner, 1986; Reys et al., 2012). At the enactive level, students build initial connections between new knowledge elements by participating in activities that involve manipulating, constructing, and arranging real-world objects related to the concept. At the iconic level, students strengthen the previously formed connections and build further connections by participating in activities involving using pictures, images, or other representations of the concept. By the time students reach the symbolic level, a schema has been formed and students are ready to take part in activities that help them connect their work at the enactive and iconic levels to abstract symbolic representations of the concept. With time, students are able to manipulate and use the symbolic representations flexibly and efficiently without referring to their enactive or iconic counterparts. This may indicate that their schema has been encapsulated as a single knowledge element.

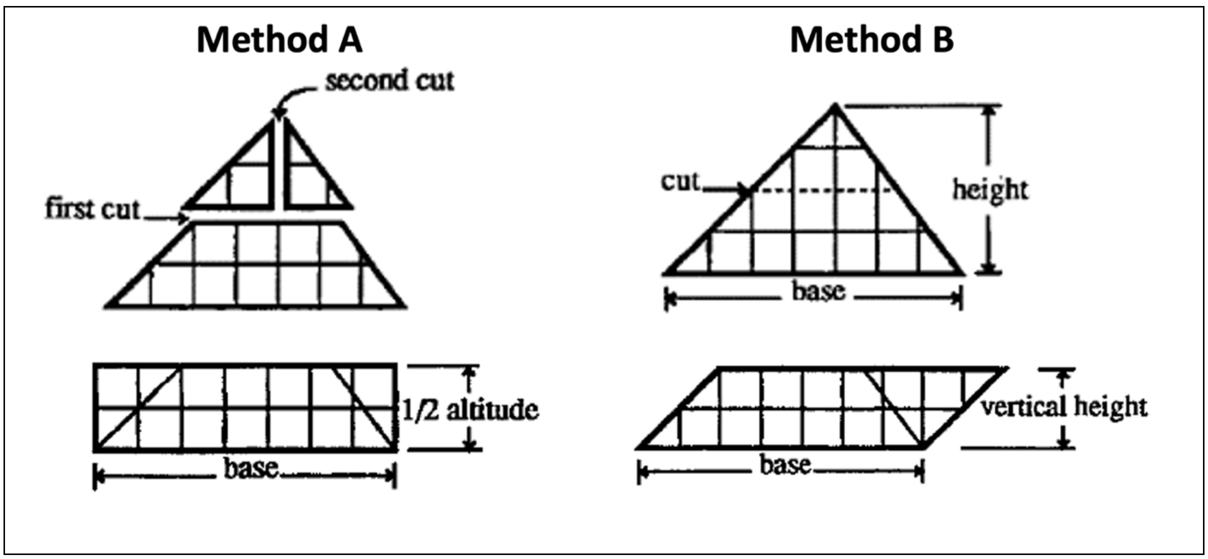
Area of Polygons

In our quest to develop a more effective lesson, we began by consulting the literature. Our search revealed an article by Neatrour (1991), which catalogued his methods for demonstrating various polygonal area formulas. We were especially interested in his two methods for the area of a triangle.

For his first method, he began by making a cut parallel to one of the bases of the triangle through the midpoint of the triangle's height (or altitude). See Figure 3 for an illustration.

Figure 3

Neatrour's (1991) Methods for the Area of a Triangle.



He then cut the resulting smaller triangle on top into two smaller right triangles. He rotated each smaller right triangle by 180º and translated them to the right and left of the bottom piece of the original triangle to create a rectangle. In Figure 3, we see that the bases of the original triangle and the rectangle are the same, but the height of the rectangle is now half the height of the original triangle. Thus,

$$A_{original\ triangle}=A_{rectangle}=base\times\frac{1}{2}height=\frac{1}{2}\times b\times h$$

For his second method, he began by making the same parallel cut as before. However, this time, he simply rotated the resulting smaller top triangle 180º and translated it down to create a parallelogram. This time, if a student knows the parallelogram area formula, the triangle area formula is forthcoming:

$$A_{original\ triangle}=A_{parallelogram}=base\times\frac{1}{2}height=\frac{1}{2}\times b\times h$$

We noticed that Neatrour's (1991) methods focused both on conserving area and halving attributes while at the same time, involved grid squares. While we appreciated his approach, we felt our students might become overloaded if these intrinsic knowledge elements were introduced initially and simultaneously. Instead, we wanted students to have plenty of room in working memory to grapple with the intrinsic idea that a triangle is half a parallelogram. This was the overall relationship we wanted them to experience and remember.

Methodology

This study builds on an existing study utilizing a qualitative, grounded theory, pedagogical action research design (Norton, 2018), in which MTE faculty collaborated to revise the guided notes for mathematics content courses for PSTs. The MTEs in the original study utilized this approach as we explored the pedagogical issue of students' struggles with mathematics vocabulary and methodically developed steps to address these issues. This work resulted in the development of the GMaV Framework detailed earlier (Bullock et al., 2021, see Figure 2) and led to the development of assessment tools for mathematics vocabulary (Ray et al., 2023).

Research Question

In our current study, the authors, two MTEs from the original research team, extended the existing pedagogical action research efforts, to address the pedagogical issue of Math 1385 students' struggles related to polygonal area. Building on the larger group's curriculum revision efforts, we sought to find research-based ways to refine and adjust our existing polygonal area lesson to address these struggles. Thus, we asked the following research question: How can we utilize cognitive load theory and Bruner's levels of developmental learning (Ayres, 2006; Pass et al., 2003) to adapt an existing lesson to address students' struggles specifically related to triangular area?

Data Collection and Analysis

As part of the larger study, during the Spring 2022 semester, we began reviewing the 24 sets of guided notes from Math 1385, along with another MTE from the research team. From this, we decided the guided notes for the "Area of Polygons" lesson would be a beneficial candidate for revisions. The lesson revisions would be our main source of data. This decision came in response to observed students' struggles with polygonal area lingering after the original lesson was taught.

To analyze the lesson revisions, we conducted iterative, thematic analyses and case comparison (Corbin & Strauss, 2008; Glesne, 2006) during the Fall 2022 semester. Here, we analyzed the iterations of the lessons based on the content and nature of the lessons. Then, we compared these iterations to one another and to the polygonal area approaches found in the research literature (Neatrour, 1991). To summarize our findings, we visually represented the approaches of the polygonal area lessons in comparison to one another to explore how these approaches could be viewed through the lenses of cognitive load theory and Bruner's levels of developmental learning.

Findings: Iterations of the "Area of Polygons" Lesson

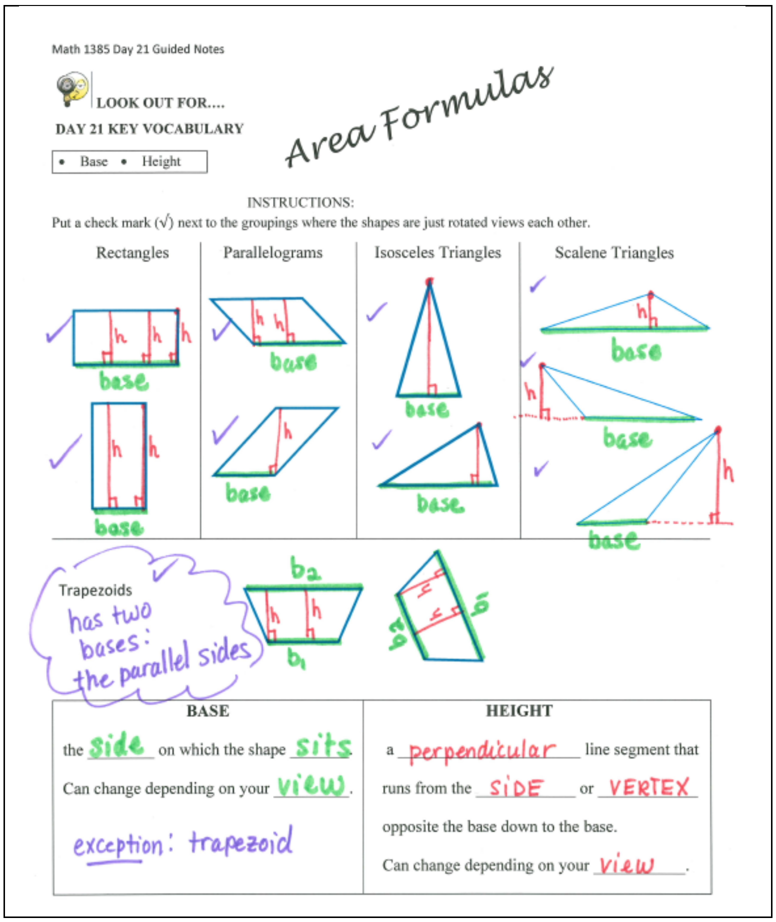
In our findings, we detail our adjustments and refinements to the lesson, specifically focused on the lesson portion involving triangular area. We outline what we found when analyzing a) the lesson portion leading up to triangular area, b) the original lesson portion involving triangular area, and c) the iterative refinements of the triangular area portion, as informed by cognitive load theory and Bruner's levels of developmental learning.

Background: Leading Up to Triangles

In this study, our focus is on our revisions of the triangular area portion of the lesson. However, during the lesson's initial portions leading up to triangle area, students explored pertinent vocabulary and other area formulas intrinsic to the triangle area formula. Students began by exploring the definitions of base and height and labeling the bases and heights of various sets of congruent shapes in different orientations. See Figure 4 for this information.

Figure 4

Exploring Bases and Heights in Different Orientations.



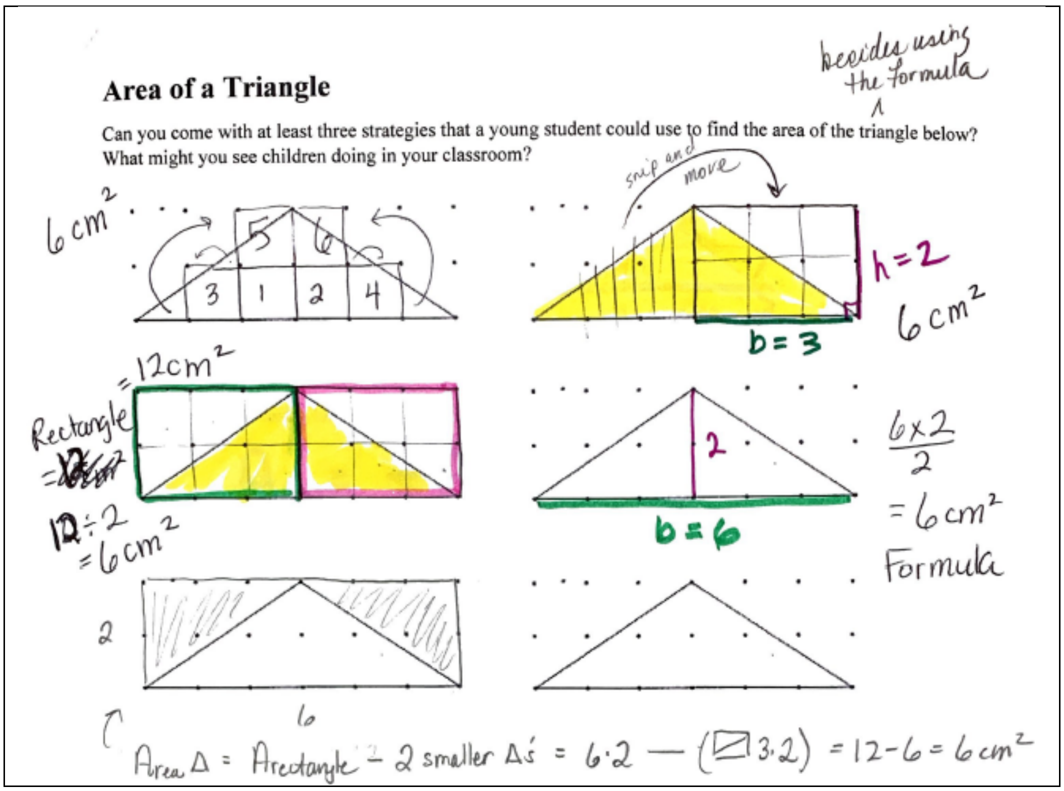
By providing both visual and verbal representations of this vocabulary at the beginning, we were carefully timing the introduction of knowledge elements intrinsic to the triangle area formula so that the terms would be available for students to use throughout the lesson. Next, the class studied the rectangle, dividing it into an array of smaller squares to better understand why its area formula is base times height. From there, the lesson moved to parallelograms where students cut parallelograms into two parts along their heights, rearranged the resulting pieces to make a rectangle, and thus, showed that area is still just base times height. Here again, we were scaffolding the recall of the intrinsic elements of area and parallelogram in preparation for the triangle area formula.

Cory's Original Triangle Area Approach (Version 1.0)

In the original triangle area portion of the lesson, Cory began by giving students six congruent isosceles triangles in their guided notes. She asked student to come up with at least three strategies a young child might use to find the area of the triangle. In the larger class discussion, students shared multiple different approaches. See Figure 5 for details on these approaches.

Figure 5

Student Strategies for the Area of a Triangle (Version 1.0).

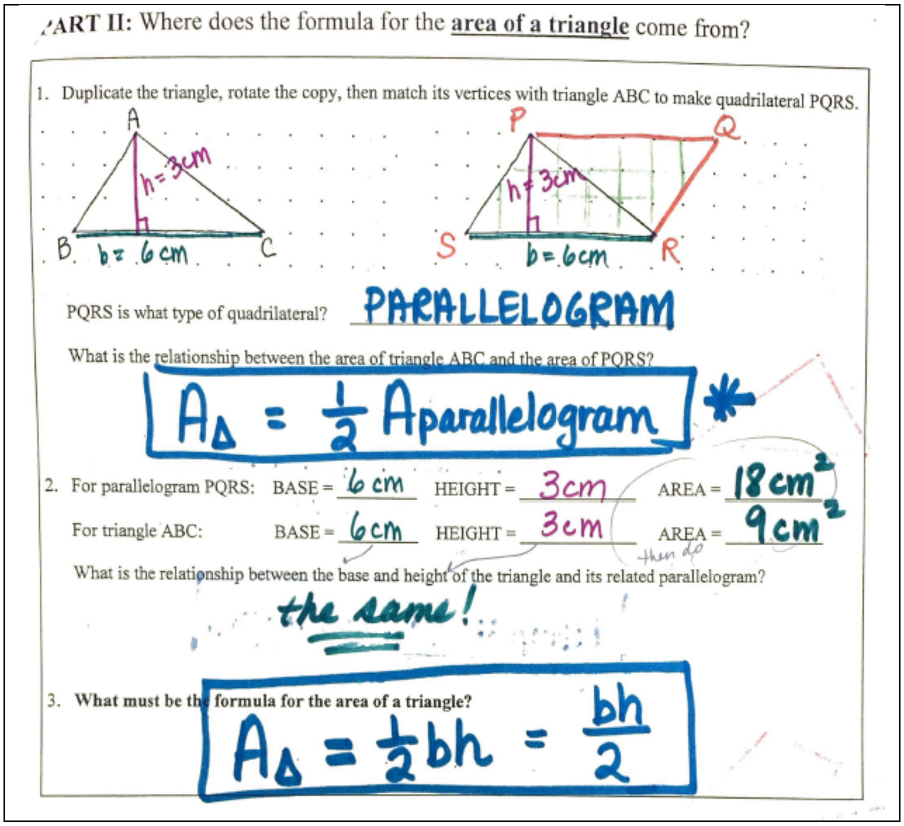


One student split the triangle up into small squares, piecing together left-over parts to make full squares. Another student split the triangle into two smaller right triangles and put the two right triangles together to make a rectangle with the same area as the original triangle. A third student also split the triangle into two right triangles but copied the two right triangles to make a large rectangle with area double the original triangle. Surprisingly, a fourth student found the area of a large rectangle encompassing the triangle. She then combined the extra areas inside the large rectangle but outside the triangle to form a smaller rectangle and subtracted the area of the smaller rectangle from the area of the large one. A fifth student was determined to use the triangle area formula, even though Cory had requested them not to.

After this exploration, Cory narrowed the focus to one specific method for finding the area of a triangle. This method is illustrated in Figure 6.

Figure 6

Original Triangle Area Approach (Version 1.0).



Students copied a given triangle on a patty paper, rotated it upside down, and traced it next to the original triangle to make a parallelogram. The class discussed the fact that the area of the triangle is one-half the area of the parallelogram. Students also noticed that the base and height of each triangle and resulting parallelogram are the same. Therefore, the class concluded that the formula for the area of the triangle must be one-half the formula for the area of a parallelogram (i.e., $A_{Triangle} = \frac{1}{2} \times b \times h$). Cory then had the class practice this method on another triangle. However, as detailed above in Kayla's approach, many students did not seem to retain the importance of the $\frac{1}{2}$ in the area of a triangle formula.

After discussing this original version of the lesson, we realized that if we wanted students to grasp this concept, we needed to eliminate any extraneous cognitive load and decrease any intrinsic load, which we considered unnecessary at this point in the learning process. While not useless, the first activity with the six isosceles triangles, as well as the emphasis on conserving the base and height in the second activity, could be deemed either extraneous or not intrinsic to the overarching importance of the $\frac{1}{2}$ in the triangle area formula. We knew we needed to revise this lesson with the goal in mind.

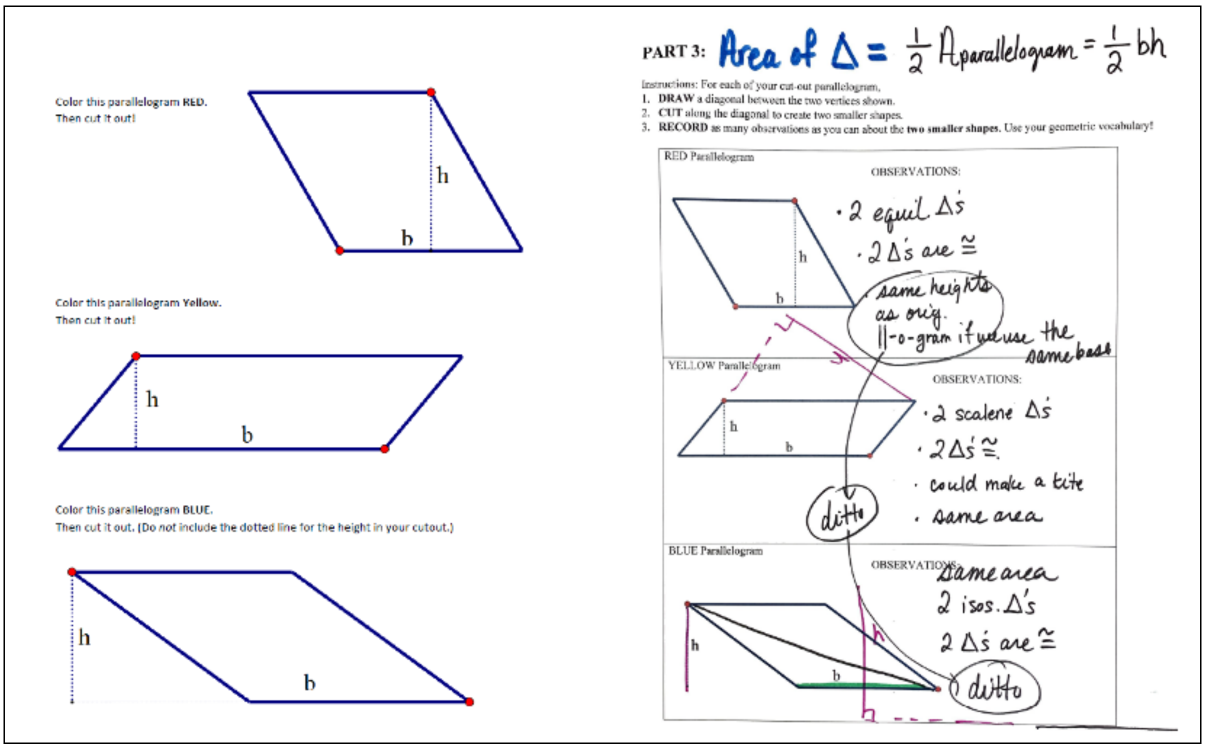
Cory's Approach (Version 2.0)

For our first attempt at a new lesson on triangle area, we created three different parallelograms, labeled the base and height in each one, and strategically marked two opposite vertices. See Figure 7 for this information. The students were to color and cut out the three parallelograms before class began. During the lesson, students followed the instructions below in their small groups:

1. **DRAW** a diagonal between the two vertices shown.
2. **CUT** along the diagonal to create two smaller shapes.
3. **RECORD** as many observations as you can about the two smaller shapes. Use your geometric vocabulary!

Figure 7

Cory's Approach (Version 2.0).



Cory then facilitated a whole class discussion about students' observations. For each parallelogram, as anticipated, the students observed that the two smaller shapes were congruent triangles with the same area. They also named the two triangles with the appropriate descriptors (equilateral, scalene, and isosceles). Cory also hoped that the strategic marking of the vertices might help the students easily notice that the base and height of each parallelogram, and its resulting triangles, are the same without causing an intrinsic overload. However, only one student mentioned this. Furthermore, she commented that it was difficult to see the height on the third parallelogram since it had to be drawn outside the shape. The discussion ended by the class highlighting the relationship between the area of a triangle and the area of a parallelogram: No matter the triangle type, a triangle is always half a parallelogram. Therefore, the triangle area formula is:

$$A_{triangle} = \frac{1}{2}A_{parallelogram} = \frac{1}{2} \times base \times height = \frac{1}{2} \times b \times h$$

We felt this lesson was an improvement over the previous version because it did a better job highlighting the structural relationships between the areas of a triangle and a parallelogram without overloading the students with untimely intrinsic elements. However, although the visuals seemed powerful, many students did not appear to deeply engage with the concept. We pondered: Had the students developed robust, meaningful, and lasting connections between triangles and parallelograms? Had their germane load been increased enough for students to begin construction of a triangle area schema in long-term memory?

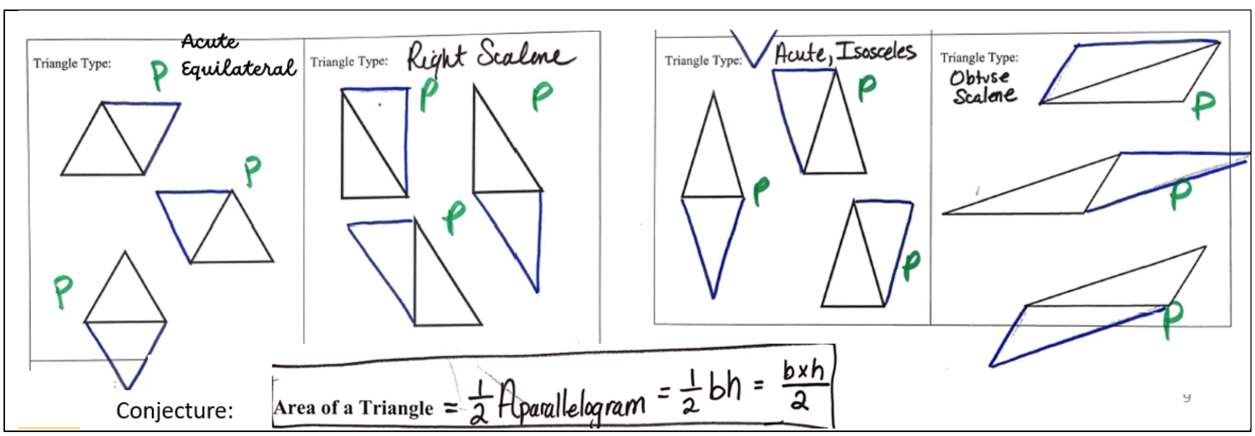
Cory's Approach (Version 3.0)

Our next reimagining of the triangle area lesson began by dividing the students into groups of four and assigning each group member one of the sets of three congruent triangles in their guided notes. Each set contained a different triangle type. Cory asked the students to do the following with their set See Figure 8 for an illustration:

1. **NAME** the triangle with two vocabulary words: Equilateral, Isosceles, Scalene, Acute, Right, Obtuse.
2. **COPY** the triangle onto patty paper.
3. **ROTATE** the patty paper triangle upside down.
4. **TRACE** the upside-down triangle next to the original triangle to a create a 4-sided shape. See if you can do this in 3 different ways!
5. **WRITE** "P" next to the 4-sided shapes that are parallelograms. Share your work with the others in your group.

Figure 8

Cory's Approach (Version 3.0)



After students completed the work for their set, Cory displayed a completed worksheet so everyone could see the correct results for all sets. As in the earlier versions, the lesson ended with a discussion emphasizing that, no matter the triangle type, the area of a triangle is half the area of a parallelogram, along with an opportunity for students to connect this concept to the symbolic formula:

$$A_{triangle} = \frac{1}{2}\, A_{parallelogram} = \frac{1}{2} \times base \times height = \frac{1}{2} \times b \times h$$

This lesson iteration provided an opportunity for students think more deeply about the different ways two congruent triangles could form a parallelogram (an increased germane load). Also, we wondered if seeing the numerous Ps covering their worksheets would make a helpful impression (an increased but timely intrinsic load). At the same time, the lesson posed some difficulties. Many students struggled to rotate their patty paper triangles 180°, creating kites rather than parallelograms (an extraneous load). We wondered what we could do reduce any confusion with the patty paper tool, yet still retain the germane problem-solving aspect of the lesson.

Cory's Approach (Version 4.0)

Version 4.0 of the triangle area lesson began by dividing the students into groups of four and giving each group member a pair of congruent laminated triangles, each pair of a different type. Cory asked the students to do the following:

1. Describe your two congruent triangles with two geometric terms: Acute, Right, Obtuse, Equilateral, Isosceles, Scalene.
2. How many different ways can you create a parallelogram with your two congruent triangles?

Before sending them off to work, Cory reviewed the definitions of parallelogram and kite, making sure students recalled the properties of these shapes. Then, after sufficient individual problem-solving time, various students shared their findings with the whole class, using their laminated triangles to demonstrate the different ways they made a parallelogram. Afterward, the students recorded their work using patty paper for their assigned triangle type on a worksheet similar to the one from Version 3.0. Finally, each group member shared their expertise by recording their work on their group members' worksheets for them. As in previous iterations, the lesson concluded with the full class highlighting the fact that a triangle is always half a parallelogram, no matter the triangle type, and by connecting this concept to the algebraic formula.

This fourth lesson iteration seemed to provide four main benefits. First, it scaffolded students' construction of a triangle area schema through Bruner's levels of developmental learning. In our lesson, students began by engaging in a hands-on activity involving laminated triangles, in line with Bruner's level one (enactive). Students next created pictorial views of their findings with patty paper, thus transferring their thinking to the pictorial representations detailed in Bruner's level two (iconic). Our lesson concluded with a discussion aimed at connecting the visual concept of a triangle being half a parallelogram to the symbolic ½ in the triangle area formula, aligned with Bruner's level three (symbolic).

Second, our fourth iteration involved a stronger germane load. Students were required to problem-solve as they thought deeply about the intrinsic definition of a parallelogram and as they utilized their spatial visualization skills to create parallelograms from triangles.

Third, rather than having Cory share the correct answers as in Version 3.0, this iteration made students responsible for their portion of the activity, requiring them to share their expertise with others. This gave them further opportunity to strengthen long-term connections within their triangle area schema.

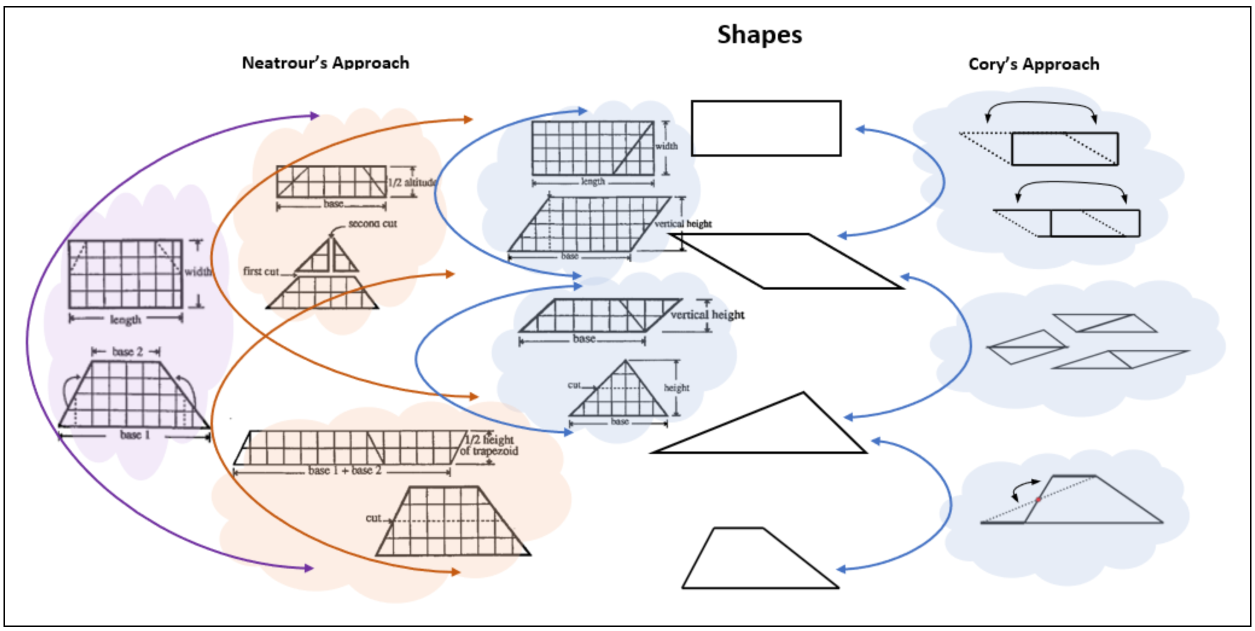
Fourth, the revised lesson focused students on the big idea that the area of the triangle is half the area of a parallelogram rather than diverting students' attention to the details of grid squares and attributes (an untimely intrinsic load), which while important, seemed to prevent some students from grasping the overall concept.

Summarizing Our Findings: Two Approaches Towards Area of Polygons

From the iterations of the triangular portion of the lesson and our comparison to Neatrour's (1991) suggested strategies for students' exploration, two main approaches towards teaching polygonal area emerged. These two approaches emphasized conservation of area of polygons in different ways. Neatrour's approach focused on the details of finding polygonal area, including figure attributes and square units, while Cory's revised approach highlighted the comparisons of areas of figures. These two approaches are summarized in the provided visual comparison show in Figure 9, where we see the strategies suggested by each approach when transforming between pairs of four different polygons – rectangle, parallelogram, triangle, and trapezoid. The arrows in the diagram indicate the two shapes involved in the corresponding transformation. Each arrow includes a visual representation of the corresponding transformation between the two indicated shapes suggested by each approach.

Figure 9

Visual Comparison of Neatrour's (1991) and Cory's Approaches.



In the visual summary (see Figure 9), Neatrour offers options for considering the relationships between area of polygons. For example, Neatrour details multiple pathways for transforming trapezoids and triangles into parallelograms or rectangles. Additionally, Neatrour's suggested strategies emphasize figure attributes, such as base and height, and how transforming shapes from one to another changes these attributes or repositions these attributes as components of a transformed shape. For example, the transformation of a triangle into a parallelogram leads to a parallelogram with the same base length as the original triangle but with a vertical height that is half the height of the original parallelogram. Additionally, looking across the suggested strategies, Neatrour's choice to impose polygons on a grid visually emphasizes square units and may encourage students to only view area as determined by the number of boxes, or units, inside the shape.

Like Neatrour, in the beginning portion of Cory's lesson, she highlighted the definitions of base and height and square units inside a rectangle. However, her revised portion of the lesson involving triangles offers an alternative emphasis. As shown in the visual summary (see Figure 9), Cory offers a single pathway for transforming between the four different shapes, from rectangle to parallelogram to triangle to trapezoid and vice versa. Essentially, these strategies build from one area formula to the next without using labeled attributes or grid squares. Thus, unlike Neatrour, Cory's revised approach includes fewer *choices* for transforming between shapes and intentionally highlights area comparison as a way for students to make connections between polygonal area formulas. Also, by refraining from labeling attributes or using grid squares, Cory's approach eliminates extra *chatter* in the form of additional detail that may be useful or meaningful for some students but not necessary to build up and explicitly connect between the polygonal area formulas. In other words, Cory's approach more readily favors the big idea of area comparison, rather than the details of the polygonal figures.

In summary, our evolving approaches towards polygonal area eliminated extra detail and provided students with a single pathway for building from one shape to the next, thus minimizing extraneous load and introducing intrinsic load in a timely fashion, rather than overwhelming working memory with multiple pathways. We posit that this approach helped students develop stronger initial connections between the polygonal area formulas. We are not suggesting that details of attributes and grid squares or the use of multiple and varied approaches are not important. Instead, we propose that more comprehensive approaches could perhaps be explored once students have a clear foundational understanding of the conceptual connections between area of polygon formulas, or that these additional approaches could allow room for differentiation when students exhibit varying levels of understanding.

Discussion and Implications

In this study, we explored iterations of an "Area of Polygons" lesson using the lens of cognitive load theory. In the process, we became aware that a thoughtful consideration of the layout of our lesson in terms of Bruner's levels of developmental learning helped us better leverage the tenets of cognitive load theory. Particularly, we were better able to scaffold students' construction of a triangle area schema by intentionally moving them through the three levels. Moreover, to increase the germane load necessary for creating stronger, stable schemas, we increased the problem-solving necessary at the enactive and iconic levels. Additionally, at each level, we recognized the need to reduce or eliminate extraneous and untimely intrinsic load which was distracting our students from grasping the big idea of the lesson. Thus, combining Bruner's levels with components of cognitive load theory provided us frameworks for revising the lesson in a powerful way.

We posit that these frameworks could be useful for effectively adapting instruction across the grade levels and mathematical content to help students create strong schemas around any mathematical topic. More broadly, we suggest that these frameworks offer flexibility in considering the unique contexts, backgrounds, and needs of learners. For example, if advanced students are ready for additional intrinsic load, they may benefit from explorations involving more detailed information that could be considered "chatter" for other students. Additionally, these frameworks provide explicit language and resources for instructors to improve curriculum materials and learning experiences across a wide range of mathematical content. Our study focused specifically on polygonal area, but we anticipate many other content areas where students traditionally struggle could also benefit from a review using our combined frameworks. In conclusion, cognitive load theory and Bruner's levels of developmental learning proved to be useful lenses for our curriculum redesign efforts.

The authors received no financial support for the research, authorship, and/or publication of this manuscript.

Beth Cory (bcory@shsu.edu) is an Associate Professor of Mathematics Education at Sam Houston State University in Huntsville, TX where she teaches mathematics content courses for elementary and middle school preservice teachers as well as master's level courses for secondary teachers and community college instructors. Her current research interests involve curriculum and assessment development with a specific focus on preservice teachers' understanding of geometry concepts, mathematical vocabulary development through the use of guided notes, as well as Calculus I assessment design. In her free time, Dr. Cory enjoys playing violin and piano.

Amy Ray (aer066@shsu.edu) is an Assistant Professor of Mathematics Education at Sam Houston State University in Huntsville, TX where she teaches mathematics content courses for future elementary, middle grades, secondary, and post-secondary educators. Her research broadly focuses on mathematics teaching and assessment developments aimed at broadening what it means to know and do mathematics as well as demonstrate mathematics thinking. More specifically, she explores the role of student work as a curricular component, the framing of teachers as curriculum and assessment writers as they engage with curriculum materials, and the intentional ways in which advancements in curriculum can be implemented in assessment practices.

References

Aichele, D.B. & Wolfe, J. (2008). *Geometric structures: An inquiry-based approach for prospective elementary and middle school teachers*. Pearson Prentice Hall.

Ayres, P. (2006). Impact of reducing intrinsic cognitive load on learning in a mathematical domain. *Applied Cognitive Psychology, 20*, 287-298. https://doi.org/10.1002/acp.1245

Bruner, J. (1986). The course of cognitive growth. *American Psychologist, 19*(1), 1-15.

Bullock, E., Ray, A., Herron, J., & Swarthout, M. (2021). The GMaV conceptual framework: Constructing elementary pre-service teacher's mathematics vocabulary understanding through contextualized guided notes. *Investigations in Mathematics Learning, 13*(4), 287-302. https://doi.org/10.1080/19477503.2021.1985906.

Common Core State Standards Initiative (CCSSI). (2010). *Common Core State Standards for Mathematics*. National Governors Association Center for Best Practices and the Council of Chief State School Officers. http://www.corestandards.org

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research, 3rd edition*. Sage.

Glesne, C. (2006). *Becoming qualitative researchers: An introduction, 4th edition*. Pearson.

Neatrour, C. R. (1991). A strategy for discovering the formulas for finding the area of polygons. *School Science and Mathematics*, *91*(8), 362-66.

Norton, L. (2018). *Action research in teaching and learning: A practical guide to conducting pedagogical research in universities.* Routledge.

Pass, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instruction design: Recent developments. *Educational Psychologist, 39*(1), 1–4.

Ray, A., Herron, J., & Bullock E. (2023). Exploring mathematics vocabulary alignment in a future elementary teacher's trajectory: A case study. *Journal of College Reading and Learning*, DOI: 10.1080/10790195.2023.2214188

Reys, R., Lindquist, M., Lamdin, D., & Smith, N. (2012). *Helping children learn mathematics*, *10th edition*. John Wiley & Sons.

ELECTRONIC JOURNAL FOR RESEARCH
IN SCIENCE & MATHEMATICS EDUCATION
VOL. 27, NO. 3, 35-63

ICRSME
International Consortium for Research in
Science & Mathematics Education

Investigating Large-Scale, High School Mathematics Achievement Through the Lens of the Cognitive Domains

Lois George
The University of the West Indies, Mona Campus

ABSTRACT

A key component of student mathematics achievement relates to the cognitive domains. This paper examines student mathematics achievement in three cognitive domains (knowledge, comprehension, reasoning) as per different achievement grades, across four years (2015, 2016, 2018, 2019), and by gender. This study used the Caribbean Secondary Education Certificate (CSEC) mathematics results across four years from the Caribbean Examinations Council (CXC), which is the main public examination board in the Caribbean. The sample constituted 69,945 public school students from 161 secondary schools in Jamaica. The study found a regular pattern of cognitive domain performance at all grade levels and in each of the four years under consideration. Students performed best in the knowledge domain, followed by comprehension, and then reasoning. Students with the highest overall achievement demonstrated the highest achievement across the three domains and there was also a strong, significant, positive correlation between students' overall grades, that related to knowledge, comprehension, reasoning, and the cognitive domains. Another key finding was that for the knowledge and comprehension domains there was a significant difference in the performance of males and females in favour of females, but the related effect sizes were minimal. Practical implications and potential directions for future research are discussed.

Keywords: mathematics achievement, cognitive domains, gender

Introduction

In present-day societies, proficiency in mathematics is considered to be a central scholastic imperative (Mullis et al., 2012). It is a prerequisite to attaining educational and vocational success, especially in careers related to science, technology, mathematics, and engineering (STEM), and in navigating daily living (Hefty, 2015; Siegler et al., 2012). At the end of high school, evidence of students' mathematics proficiency is often the attainment of a qualification that indicates a passing grade in some standardized, high stakes, exit examination. This qualification is significant since it serves as the main matriculation requirement for further studies and employment.

Performance outcomes in the form of scores, or grades students receive from completing these mathematics examinations, reflect their knowledge and skills of content-related subject matter such as that related to specific strands in mathematics like Geometry or Algebra. For some examination bodies or examining boards, these scores also reflect cognitive dimensions or domains, such as knowing, applying, and reasoning which relate to the thinking processes that students are expected to utilize as they engage with different mathematical topics and tasks (Harks et al., 2014; Mullis et al., 2020). In regard to the cognitive domains in particular, which is a key focus of this paper, Nilsen et al. (2016) stated that students "do not just need knowledge in mathematics, but must also be able to apply knowledge and conceptual understanding in different contexts, and to analyze and reason to solve problems" (p. 7). This is highlighted as a key outcome of mathematics teaching and learning in frameworks in mathematics (e.g., Schoenfeld & Kilpatrick, 2008), mathematics curricular/policy documents (Caribbean Examinations Council, 2015; Department for Education, 2013; National Council of Teachers of Mathematics, 2014) and international assessment frameworks in mathematics such as the Trends in International Mathematics and Science Study (TIMMS) and Programme for International Student Assessment (PISA) (Mullis et al., 2020; *Programme for International Student Assessment*, 2021).

While students' scores capture multiple elements of their knowledge and skills, Steiner and Ashcraft (2012) note that this achievement is not uniform. In other words, two students who obtain the same grade in an assessment may differ widely in their cognitive dimensions and/or content knowledge. Dogan and Tatsuoka (2008) also note that assessment reports that provide only total scores of examinees are limited in that they do not offer insights into areas of students' mathematics difficulties. Consequently, the present author asserts that investigating achievement from the perspective of overall grades provides just one of several possible dimensions related to student achievement. It is important that research focus on a multiplicity of perspectives in order to gain a more in-depth and comprehensive portrait of student achievement. Further, unpacking students' achievement in high stakes examinations also provides insightful observations to countries, in particular "educational policymakers, administrators, teachers, and researchers …in understanding educational outcomes more fully, which is a core concern of effective educational planning and reform" (Mullis et al., 2012, p. 1). An example of this is seen in research undertaken by Shalem et al. (2013) who reported that as South African teachers participated in curriculum mapping of large-scale assessments, they gained an understanding of both the specific, grade-level content that was assessed and the related cognitive demand. Teachers were then better able to evaluate their classroom practice and understand the curriculum.

While there are several frameworks for presenting elements of the cognitive domain, arguably, Bloom's taxonomy is one of the most widely used and well-established in education, especially in relation to developing achievement tests and interpreting these test results (Webb, 2020). The original taxonomy, that has since been revised, is organised within a triangle and comprises six levels related to the cognitive domain: knowledge, comprehension, application, analysis, synthesis, and evaluation. The present author notes that although extensively adopted in practice, Bloom's taxonomy has also faced considerable critique. For example, whereas many practitioners perceive the elements of the taxonomy to be hierarchical (Webb, 2012), this notion has been disputed and instead it has been proposed that these components are merely a set of categories (Postlethwaite, 1994). Notwithstanding the aforementioned, the elements of Bloom's taxonomy build on each other (Thomson, 2006). Mullis et al. (2003) also alluded to this and stated that "facility in using mathematics, or reasoning about mathematical situations, depends primarily on mathematical knowledge" (p. 27). Mullis and Martin (2017) add that "without access to this knowledge base … students would find purposeful mathematical thinking impossible" (p. 23). A comprehensive critique of Bloom's taxonomy is beyond the scope of this paper. However, Webb (2020) notes that although not explicitly acknowledged, the influence of this taxonomy on large-scale, high stakes assessments, such as the TIMMS is apparent.

To date, students' cognitive dimensions (e.g., knowing, understanding, reasoning) in examinations have received far less scholastic attention than content-specific components of a given curriculum (Harks et al., 2014). The present author surmises that this may be related to the performance outcome reporting formats that examining bodies use, and/or the type of data these bodies collect on student achievement. For example, very few examining bodies report student outcomes as per cognitive dimensions. Thomson (2006) also notes that

> The content domains are fairly, consistently, and readily found in the curricula of the participating countries, and are the subject of the major international and national reports for

> TIMSS. Developing reliable and valid achievement scales for the cognitive domains is not as straightforward. (p. v)

Notwithstanding the aforementioned, since the development of a range of cognitive skills in mathematics is an educational imperative, and that there is general consensus within mathematics education that assessing and reporting elements of the cognitive domain is important (Harks et al., 2014), research, such as the current one, that focuses on exploring student achievement in this regard is significant. In addition to this, Suurtamm et al. (2016) note that "If the enacted curriculum of the classroom and the assessed curriculum are to inform each other and to enhance student learning in positive and productive ways, then large-scale, external assessments cannot operate in isolation from the classroom" (p. 22). This assertion points to another benefit of this research in relation to its potential to positively impact teachers' classroom practice.

Research Context: Assessing and Reporting Mathematics Achievement

CSEC Mathematics Examination

In the English speaking Caribbean, at the end of high school, most students generally sit for the Caribbean Secondary Education Certificate (CSEC) mathematics examinations to obtain their qualifications in mathematics. These examinations are administered by the examining body, the Caribbean Examinations Council (CXC[1]), and are offered at the General Proficiency level in June and January annually. In terms of international comparison, the CSEC mathematics syllabus states that "the competencies and certification acquired upon completion of this course of study are comparable with the mathematics curricula of high schools world-wide" (Caribbean Examinations Council, 2015, p. 1). Hence, the present author opines that the findings of this study are very likely to be important and applicable beyond the local context.

The CSEC mathematics examination comprises an external and internal assessment component weighted 80% and 20%, respectively. The external component consists of two papers, Paper 01 and 02, that account for 80% of the overall grade. Paper 01 is worth 30% of the final grade and has 60 compulsory, multiple-choice items while Paper 02 includes ten, compulsory, constructed-response items, based on the nine topics and associated objectives covered in the mathematics syllabus. Paper 02 contributes to 50% of the final grade. The internal assessment (Paper 03) is school-based and requires examinees to demonstrate the application of mathematics in real life situations. In this regard it comprises a project on any relevant topic, or combination of topics, that is assessed internally by the teacher and externally by CXC (Caribbean Examinations Council, 2015).

CSEC Grade Reporting

For the CSEC mathematics qualification, student achievement is reported as grades denoted by Roman numerals ranging from I-VI with I to III designated as passing grades. Each grade band within each grading system has a cut-off point, however, CXC does not make its cut-scores public (McPherson, 2020). Therefore, the range of raw scores within each grade is not known. Also, the

---

[1] The Caribbean Examinations Council (CXC) is an examination board that conducts examinations, and awards certificates and diplomas based on the results of these examinations in 16 Caribbean countries. The Caribbean Secondary Education Certification (CSEC) is one of several qualifications offered for a range of subjects that individuals would enrol in in high school. The CSEC typically and currently serves as the main matriculation qualification for entry into postsecondary education in the Caribbean.

percentage of the test that students would need to get correct to be considered passing is also not made public. Table 1 presents the grades and the associated descriptors of performance.

Table 1

Descriptors of CSEC Performance Outcomes

Overall Grade	Performance Descriptor	Profile Grade	Performance Descriptor
I	Outstanding	A	Outstanding
II	Good	B	Good
III	Fairly good	C	Fairly good
IV	Moderate	D	Moderate
V	Limited	E	Weak
VI	Very limited	F	Poor

For the CXC, beyond reporting grades as Roman numerals, it reports achievement relating to the cognitive domain as three *profile dimensions* on a scale of A-F (see Table 1). The profile dimensions include knowledge (P1), comprehension (P2), and reasoning (P3), and specify the cognitive demand of the items and questions in the CSEC mathematics examination (Caribbean Examinations Council, 2015). They also capture the thinking processes expected of students as they engage with the mathematics content and are similar to some of the categories in the cognitive domain included in Bloom's Taxonomy (Bloom et al., 1956). For the CSEC mathematics examination, 30% of the items are allocated to knowledge and reasoning, respectively, while 40% are aligned to the comprehension dimension. Table 2 provides a breakdown of marks and percentage weightings in the examination components by profiles.

Table 2

Assessment Items by Profiles and Item Format on the CSEC Mathematics Examinations (Caribbean Examinations Council, 2015, p. 5)

Profiles	No. of Marks in Examination Components			Total
	Paper 01 Multiple Choice	Paper 02 Constructed Response	Paper 03 School Based	
Knowledge	18	30	12	60 (30%)
Comprehension	24	40	16	80 (40%)
Reasoning	18	30	12	60 (30%)
Total	60 (30%)	100 (50%)	40 (20%)	200 (100%)

The Caribbean Examinations Council (2015) informs that the profile '*Knowledge*' requires examinees to recall rules, procedures, definitions, and facts. Items that align to this profile are characterised by rote memory, simple computations, and constructions, while '*Comprehension*' necessitates algorithmic thinking, whereby algorithms are used and applied to familiar problems. In this context, therefore, comprehension is not considered to be the same as reasoning and solving non-routine problems. It further adds that the '*Reasoning*' profile dimension encapsulates several competencies. These include the (i) translation of non-routine problems into mathematical symbols and then choosing suitable algorithms to solve the problems; (ii) combination of two or more algorithms to solve problems; (iii) use of an algorithm or part of an algorithm, in a reverse order, to solve a problem; (iv) making of inferences and generalisations from given data; (v) justification of results or statements; and (vi) analysis and synthesis of mathematical data. Two exemplars of specimen CSEC mathematics examination questions, the related solutions, and related profile dimension allocation are presented in Figures 1-4 in the Appendix.

The profile dimensions of the CXC align very closely with the notion of cognitive domains or dimensions used by the TIMSS, which is an international assessment of mathematics and science at Grades 4 and 8. TIMMS has been conducted quadrennially since 1995. Similar to the CXC, the TIMMS is large-scaled and reports student achievement using three cognitive domains, namely, knowing, applying and reasoning. Mullis and Martin (2017) inform that:

> knowing, covers the facts, concepts, and procedures required by students, applying, focuses on the ability of students to apply knowledge and conceptual understanding to solve problems or answer questions, and reasoning, goes beyond the solution of routine problems to encompass unfamiliar situations, complex contexts, and multistep problems. (p. 22)

Table 3 shows the distribution of assessment items by cognitive domain and item format.

Table 3

Distribution of Assessment Items by Cognitive Domain and Item Format in the 2011 TIMMS

Cognitive Domain	Multiple Choice	Constructed Response	% of Score Points
Knowing	43	30	39
Applying	34	41	41
Reasoning	16	21	20
Total	93 (50%)	92 (50%)	100

The TIMMS assessment differs from that of the CSEC in that it is conducted with Grade 8 students, while the CSEC examination is typically completed by students at the end of high school. Additionally, the TIMMS assesses four content domains- Number, Algebra, Geometry, Data and Probability, whereas for the CSEC, nine content domains are assessed. In this paper cognitive domains/dimensions and profiles are used interchangeably.

Overview of CSEC Mathematics Achievement in Jamaica

To provide a context for examining Jamaican public school students' mathematics achievement in each of the cognitive domains, this section provides an overview of their overall mathematics achievement from 2009-2019 (see Table 4) in the CSEC mathematics examination which is the main exit mathematics examination completed by Jamaican students at the end of high school.

Table 4 shows that over the period of 11 years from 2009-2019, on average, slightly over 50% of the students who wrote the examination did not obtain a passing grade. This suggests that after more than 11 years of formal, compulsory mathematics instruction, many students lacked the requisite mathematical knowledge and skills to function effectively in everyday life and to pursue jobs and educational opportunities that required a qualification in mathematics. This trend is consistent with many international jurisdictions, whose learners also struggle with mathematics (Fenanlampir et al., 2019; Nelson & Powell, 2018).

Student achievement as an educational concern and research focus is multi-dimensional and can be explored in varied ways. While the data presented in Table 4 provides a general summary of Jamaican students' mathematics achievement, as previously stated, deeper insights can be gleaned by focusing on different aspects of another layer of this achievement, the profile or cognitive dimensions. In Jamaica, this has not been the focus of previous empirical exploration, but Cato (2020) explored this among a sample of 370 students from the island of St. Vincent in the Caribbean using the 2017 May/June CSEC mathematics examination. This research differs from that of Cato (2020) in that it uses population data and includes multiple examination years. To date, research related to student mathematics achievement in Jamaica have only focused on students' overall grades (e.g., Crossfield & Bourne, 2017; George, 2020; George, 2013; Spencer-Ernandez & George, 2016).

Table 4

CSEC Mathematics Examinations Percentage Pass for Jamaican Public Schools From 2009-2019

Year	No. Sitting	Grades			Pass	
		I	II	III	No	%
2009	19,990	1,623	2,508	4,054	8,185	40.9
2010	20,742	2,029	2,876	4,366	9,271	44.7
2011	20,850	1,652	2,527	4,139	8,318	39.9
2012	23,729	1,909	2,583	4,398	8,890	37.5
2013	22,870	1,764	2,910	4,985	9,659	42.2
2014	23,351	2,955	4,015	5,993	12,963	55.5
2015	23,639	4,203	4,692	5,762	14,657	62.0
2016	23,993	3,063	3,123	5,270	11,456	47.7
2017	23,567	2,751	3,312	5,775	11,838	50.2
2018	22,214	2,793	4,705	5,347	12,845	57.8
2019	21,320	1,748	4,212	5,685	11,645	54.6
Average (%)		**10.8**	**15.2**	**22.6**		**48.5**

Note. Data from the Jamaica Ministry of Education

Another aspect of achievement that this research explores is that relating to students who are near proficiency levels or who are on the border of passing or failing high-stakes tests. This group of students have been given a variety of labels, such as bubble students or kids (McNeil, 2002), cusp children (Bradbury et al., 2021), borderline students (Amrein-Beardsley & Berliner, 2002) or near-passing students (Rothman & Henderson, 2011) and have been given considerable focus in practice (Minarechová, 2012). One reason for this emphasis is the significant impact that these students may have on a school's or district's composite scores on a high-stakes test (Minarechová, 2012). In this regard, Reback (2008) points out that in systems where examination pass rates are prioritised, borderline students would have the greatest impact on a school's performance measure. Another reason is the perception and research finding that enhanced provisions to assist this group in making gains in their achievement would allow them to pass high-stakes examinations, whereas this outcome would be unlikely for lower-performing groups (Hutchings, 2015; Marks, 2014; Reback, 2008).

Considering the deleterious impact of low or no mathematics qualifications on an individual's educational and job prospects, the focus on this borderline group is important. McMahon (2022) notes that there has been limited scholastic attention on the achievement of these students. Additionally, the research context, like many developing countries internationally utilise a minimum competency school accountability system that only include students' test scores via pass rates. Also, the pass rate on mathematics examinations has been historically low and resources are limited, therefore, there would be an interest and emphasis in improving the performance of students who are on the margin of passing by first exploring their examination performance in greater detail. McMahon (2022) adds that a focus on this research domain could provide educational stakeholders with insights that could assist them in more targeted interventions geared towards meeting the needs of this student group in a more targeted way.

In this research, students who obtained Grades III and IV would be borderline students. While the possible grades for the CSEC mathematics examinations range from I-VI, this research focuses on students who received Grades I-III, which are passing grades, and Grade IV, which is a borderline failing grade. Grades V and VI are therefore not included in the current data analysis but could be incorporated in future research. Students who obtained a borderline failing grade have not been the focus of previous research relating to cognitive dimensions in mathematics assessments.

Stemming from the aforementioned, this paper aims to add to the existing literature both globally and locally by investigating the profile dimensions in general and for the years 2015, 2016, 2018, 2019 as per students' examination grades and gender. In particular, this research addresses the following research questions:

1. (a) How does the profile (knowledge, comprehension, reasoning) performance of students with Grades I-IV compare?
   (b) What is the relationship between:
      (i) students' overall mathematics examination grade and each of the three profile grades?
      (ii) the knowledge and comprehension domains; the knowledge and reasoning domains; the comprehension and reasoning domains?
2. How do the profile grades bands (knowledge, comprehension, reasoning) of students with passing grades compare (i) by year (ii) in general?
3. How do the profile grade bands (knowledge, comprehension, reasoning) of students with a borderline failing grade (Grade IV) and a borderline passing grade (Grade III) compare?
4. Is there a statistically significant difference in the profile grades of students with passing grades by gender?
   H_o: There is no significant difference in the profile grades of males and females.

Cognitive Dimensions in General and by Gender

Since research concerning the cognitive dimensions as conceptualised in this research is limited, relevant research related mainly to the TIMMS assessment, which reports student achievement as per the cognitive domain, is reviewed in this section. For the TIMMS assessment data for 2011, 2015 and 2019, there does not appear to be a pattern in achievement related to the three cognitive domains of knowing, applying and reasoning. Based on the TIMMS 2019 mathematics assessment, Mullis et al. (2020) report that more countries had a weakness in the knowing domain than in the applying and reasoning domains. This contrasts with the 2011 assessment in which more countries demonstrated relative strengths in knowing mathematics (i.e., recalling, recognizing, and computing) than in applying mathematical knowledge and reasoning (Mullis et al., 2012). For the 2019 sitting, reasoning was reported to be a relative weakness for approximately 44% (n = 28) of the countries while applying was reported to be a relative strength for about 38% (n = 24). Mullis et al. (2012) informed that generally, the TIMSS 2011 participants with the highest achievement overall also had the highest achievement across the cognitive domains.

Kaleli-Yılmaz and Hanci (2016) investigated the cognitive domain components of student mathematics achievement in general and by gender of 652 eighth grade Turkish students (305 girls and 347 boys). They used items from the TIMMS 2011 mathematics test and reported that students performed best and worst on the applying and reasoning domains, respectively. Thomson (2006) who explored TIMMS 2003 data with a specific focus on Australian student performance found that for both years four and eight, students performed best in the reasoning domain. Furthermore, for Grades four and eight, respectively, achievement relating to the applying and knowledge domains were lower compared to the other domains. As it relates to the relationship between the cognitive domains and mathematics achievement, Pogoy et al. (2015) used TIMMS 2011 data across countries and found a large, significant, positive relationship between each of the three cognitive domains and mathematics achievement.

The present research aims to add to this existing empirical data by exploring assessment data related to cognitive domains across four years instead of quadrennially like the TIMMS or cross-sectionally (e.g., Kaleli-Yılmaz & Hanci, 2016) in order to investigate whether there are patterns of performance. This study also explores the achievement of students with passing and borderline fail grades to investigate the cognitive domain outcomes of students with different levels of mathematics achievement to glean insights from this analysis.

As it relates to gender differences related to the cognitive domain components of student mathematics achievement, Mullis et al. (2016) found, based on the 2015 TIMMS assessment, that the assessment results "show an advantage for girls in the Reasoning domain" (p. 125). In the 2019 assessment, "boys had higher average achievement than girls in many countries in the cognitive domains—31 countries in the knowing domain, 15 in the applying domain, and 28 in the reasoning domain. Girls had higher average achievement than boys in all three domains in Oman, the Philippines and South Africa (fifth grade)" (Mullis et al., 2020, p. 77). Mullis and her associates, however, did not indicate whether these differences were significant. Kaleli-Yılmaz and Hanci (2016) also explored the cognitive domain components of student mathematics achievement by gender and found that although girls outperformed their male peers on all cognitive domains (knowing, applying and reasoning), these differences were not statistically significant. This finding from Kaleli-Yilmaz and her associate aligns with research on general mathematics achievement and gender among high school students which have found no significant gender differences (Cimpian et al., 2016; George, 2022; Lindberg et al., 2010; Lubienski & Pinheiro, 2020). Exploring achievement as it relates to gender continues to be an important component of educational research considering the variability of findings to date (Else-Quest et al., 2010; Forgasz, 2012; Leder, 2012). The sample from Kaleli-Yılmaz and Hanci's (2016) study is from Turkey, and Jamaica has not yet participated in the TIMSS. Therefore, this research endeavours to add findings from another jurisdiction, the Caribbean, related to the important research focus of the cognitive domains.

Research Design and Methods

Data for the Study

The author obtained data on Jamaican public-school students' performance on the June sitting of the CSEC mathematics examination from 2015 to 2019 from the Ministry of Education [Jamaica]. However, the profile dimension data were not available for 2017 and so were not included in the analysis reported in this paper. The deidentified data for the June offering of the CSEC mathematics examinations were chosen for analysis because most candidates in high school complete their examinations within this period. Jamaica was chosen as the focus of this study because in the Anglo-Caribbean, it has the largest number of candidates participating in CSEC mathematics examinations annually which accounts for approximately 50% of the candidates examined. This substantial sample would be adequate in providing answers to the study's research questions and could also form the basis for generalising the findings to Jamaica and the wider Caribbean. Additionally, this study builds on recent previous research conducted in Jamaica relating to mathematics achievement (George, 2022; Spencer-Ernandez & George, 2016) in order to establish a research base on this critical research topic.

Public-school student data were used because it accounts for most of the students who sit for the CSEC mathematics examinations at the end of high school and was accessible to the author through the Ministry of Education [Jamaica]. Data from private institutions, however, were not available to the author. As per the data received, 168 schools entered candidates to sit for the CSEC mathematics examination. Of the 168 schools, seven schools did not have data for all of the years

(2015, 2016, 2018, 2019) that were relevant to this study. The data for these seven schools, which corresponded to 786 students, were removed. Therefore, the data analysis centred on the data from 161 schools. The 161 schools accounted for 95.8% of the performance data for public school students for four years 2015, 2016, 2018 and 2019.

The Performance Criterion

As previously stated, the CXC reports grades from the CSEC mathematics examinations on a scale of I to VI. However, grades I – III are considered to be passing grades. The CXC also reports profile dimensions (knowledge, comprehension, and reasoning) on a scale of A-F (see Table 1).

The Sample

The research sample was taken from candidates (N = 89,719) from 161 public educational institutions who sat for the CSEC mathematics examinations in Jamaica in 2015, 2016, 2018 and 2019. For two of the four research questions (Questions two and four), data from students who passed the CSEC mathematics examinations by obtaining Grades I-III were used. This represented 56% of the students who sat for the CSEC mathematics examinations for the four years being considered in this study. For research question three, the sample comprised 21,842 and 19,700 students who obtained a Grade III (borderline pass) and Grade IV (borderline fail), respectively. Table 5 shows the number of students in and gender of the sample for each of the four years under consideration.

Table 5

Sample Demographics

Gender	Year				Total	
	2015	2016	2018	2019	No	%
Students who obtained grade I-III						
F	8,639	6,684	7,570	6,747	29,640	33.0
M	5,988	4,761	5,182	4,674	20,605	23.0
Total	14,627	11,445	12,752	11,421	50,245	56.0
Students who obtained grade IV						
F	2,032	2,146	3,690	3,474	11,342	12.6
M	1,570	1,604	2,696	2,488	8,358	9.3
Total	3,602	3,750	6,386	5,962	19,700	22.0

Data Analysis

Microsoft EXCEL 2016 and the Statistical Package for the Social Sciences (SPSS 21) were used for the data analysis for this research. Each Roman numeral and letter grade was first converted to a number (see Table 6) and then the requisite statistical analyses, as per the different research questions, were undertaken.

Table 6

Grade and Profile Transformation Summary

Grade	Equivalent	Profile	Equivalent
VI	1	F	1
V	2	E	2
IV	3	D	3
III	4	C	4
II	5	B	5
I	6	A	6

The transformation of Roman numeral grades to a numeric form has been adopted in previous research for investigating similar data from CXC (e.g., George, 2020; Griffith, 2013). While the use of large-grain data (grades/profiles) may be seen as a limitation, the unavailability of students' raw scores made this the most pragmatic approach to the data analysis. In undertaking the analysis, the author considered two options in interpreting the profile grades. The profile data could be considered to be continuous in the same way that data from a Likert scale is widely treated although this is debated (Carifio & Perla, 2008; Wu & Leung, 2017) and so parametric/non-parametric tests could be applied. It could also be conceptualised as data that is ranked, and therefore can be considered to be interval data or nominal data, and so Chi-Square tests would be performed. The author performed the analyses related to each of the previously mentioned considerations and found that for each research question, the conclusion for the statistical analyses done was the same. To answer research questions one(a), two, and three, the current author found frequencies (totals and percentages) and carried out descriptive statistical analyses (means, modes, and standard deviations). For research question one(b), a Spearman's rank-order correlation was performed, while for the fourth research question, an Independent Samples *t*-test was carried out.

In organizing and reporting the results to research questions two and three which focused on profile dimension grade bands, after first examining the profiles individually and across the four years, the number of students who obtained profiles A and B were combined to form a profile grade band for top performance. The number of students who obtained profiles C and D, as well as E and F was combined individually, to form a profile grade band for satisfactory and unsatisfactory performance, respectively. This approach allowed for more efficient reporting of results since there are six profiles and grades.

Results

Research Question 1a: How does the profile (knowledge, comprehension, reasoning) performance of students with Grades I-IV compare?

Table 7 presents summary statistics (mean, mode, and standard deviation) for the individual grades I-IV. An inspection of Table 7 shows that the students with the highest overall achievement (Grades I and II) also demonstrate through the means, the highest achievement across the three profiles.

Table 7

Summary Statistics for Students with Grades I-IV

Summary Statistics	Knowledge	Comprehension	Reasoning
Overall for Students with Passing Grades			
N	50,245	50,245	50,245
Mean	5.16	4.70	4.36
Mode	5.00	4.00	4.00
Std. Deviation	0.76	0.90	1.01
Grade IV			
N	19,700	19,700	19,700
Mean	3.70	2.90	2.48
Mode	4.00	3.00	2.00
Std. Deviation	0.61	0.55	0.56
Grade III			
N	21,842	21,842	21,842
Mean	4.53	3.88	3.53
Mode	5.00	4.00	4.00
Std. Deviation	0.54	0.47	0.52
Grade II			
N	16,623	16,623	16,623
Mean	5.40	4.91	4.50
Mode	5.00	5.00	5.00
Std. Deviation	0.52	0.39	0.58
Grade I			
N	11,780	11,780	11,780
Mean	5.99	5.91	5.72
Mode	6.00	6.00	6.00
Std. Deviation	0.10	0.29	0.46

For students who obtain the highest grade possible, which is a Grade I, the most common grade for each of the three profiles is an A (Mode = 6), while for students who obtain a Grade II it is a B (Mode = 5). The result is more nuanced for students who perform less well by obtaining a Grade III or IV. In relation to the former group, for the knowledge profile, the most common profile grade is B, while for comprehension and reasoning it is a C. For students who attain a Grade IV, only the knowledge profile has a mode of 4 (profile grade C) which represents a fairly good performance. For the comprehension and reasoning profiles, the performance is moderate and weak (profile grades D and E), respectively. For students with passing grades (Grades I-III), the modal descriptor for the knowledge profile is good (Mode = 5) while for comprehension and reasoning it is fairly good (Mode = 4). This suggests that students that pass the CSEC mathematics examination generally perform very well as it relates to the three profile or cognitive dimensions.

In addition, for all grade levels (I-IV), the knowledge and reasoning profiles have the highest and lowest means, respectively. This suggests that students generally perform best in the knowledge profile, followed by comprehension and then reasoning, in descending order. For the standard deviation, this observation is reversed. This means that for each of the three profiles, the students who obtain Grades I and IV appeared to be the most and least homogeneous, respectively, in terms of achievement. The standard deviation for knowledge, comprehension, and reasoning, respectively (0.76, 0.90, 1.01) suggest that there is greater homogeneity in profile performance for the knowledge profile in comparison to that of comprehension and reasoning.

Research Question 1bi: What is the relationship between students' overall mathematics examination grade and each of the three profile grades?

For students with passing grades (Grades I- III), a Spearman's rank-order correlation was run to determine the relationship between students' overall grade and their profile grade. Results of the Spearman correlation indicated that there is a strong, positive correlation between the overall and profile grade for knowledge, comprehension, and reasoning, which were all statistically significant (rs(50,243) = .802, p <= .001); (rs(50,243) = .900, p <= .001); (rs(50,243) = .839, p <= .001), respectively.

Research Question 1bii: What is the relationship between the knowledge and comprehension domains; the knowledge and reasoning domains; the comprehension and reasoning domains?

The relationship between profiles was also explored. The analyses found that there was a strong, positive, statistically significant correlation between the profile grades (knowledge and comprehension; comprehension and applying; comprehension and reasoning. The knowledge profile was more strongly correlated to the comprehension profile than the reasoning profile (rs(50,243) = .739, p <= .001); (rs(50,243) = .680, p <= .001), respectively. The comprehension profile was more strongly correlated to the reasoning profile than that of knowledge (rs(50,243) = .764, p <= .001); (rs(50,243) = .739, p <= .001), respectively.

Research Question 2: How do the profile grades bands of students with passing grades compare (i) by year (ii) in general?

Figure 5 presents the profile (knowledge, comprehension, reasoning) grade bands (A-B, C-D, E-F, as applicable) for students who obtained passing grades of I, II, and III across each of the years 2015 - 2019. Every student receives a grade that ranges from A-E for each of the three profile dimensions. The profile grade bands A-B, C-D and E-F indicate top, satisfactory, and unsatisfactory performance, respectively. Therefore, for 2019 for example, for all students who passed the CSEC mathematics examination for the Knowledge profile 66.5% of the cohort obtained a profile grade of A or B, while 33.5% received a grade within the C-D band. For the Comprehension profile, 51% of the students who obtained Grades I-III, received a grade within the A-B band whereas 49% of these students received a profile grade of C or D.

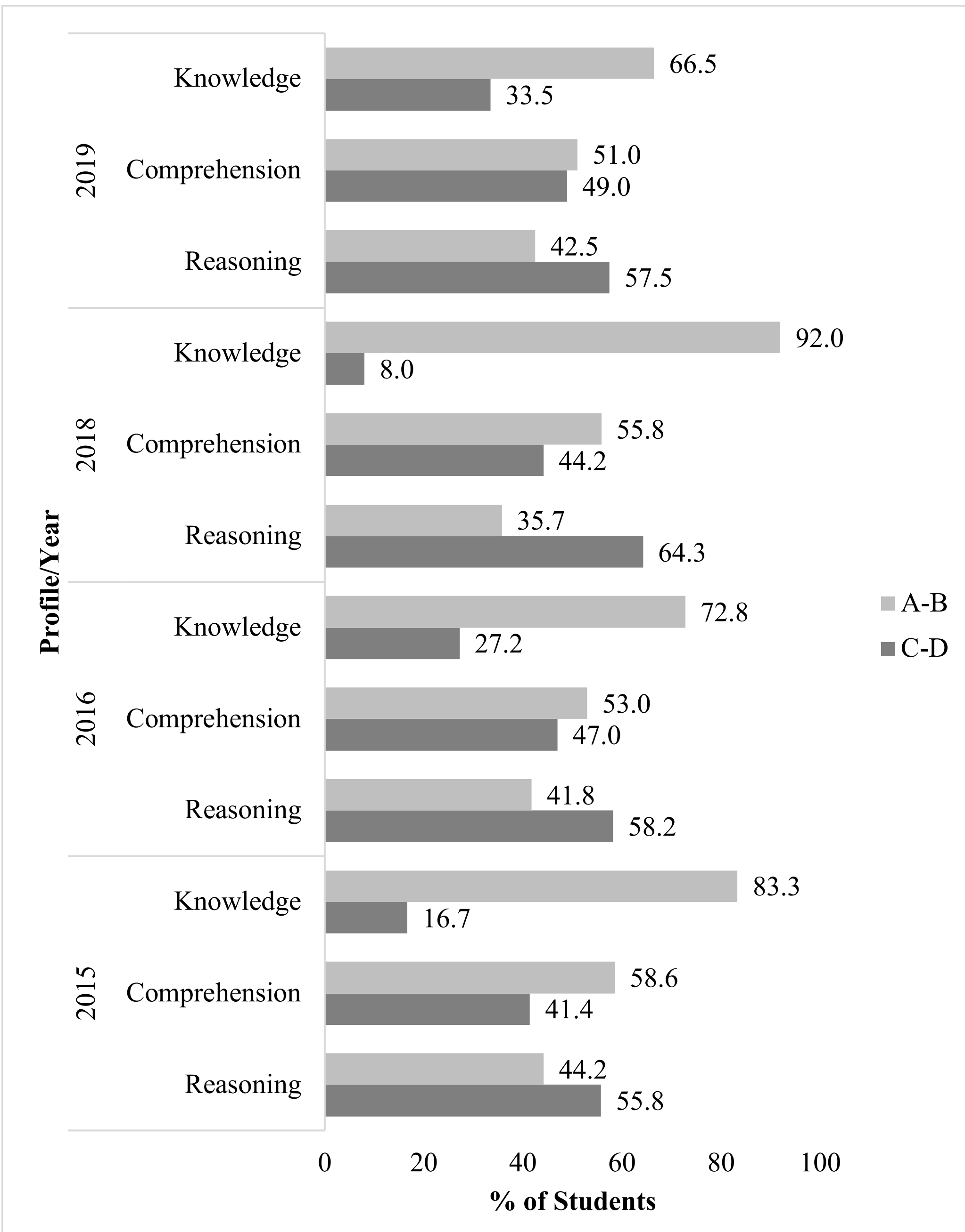
Figure 5 shows that over the years 2015, 2016, 2018 and 2019, there appears to be a regular pattern of grade band attainment as it relates to each of the profiles. For each of the four years under consideration, a close examination of the top band performance for each profile reveals that the knowledge profile has the highest percentage of students obtaining A and B grades, followed by the comprehension profile and then the reasoning profile with the smallest proportion of students exhibiting top band performance. This trend is reversed for students who performed satisfactorily by attaining profile grades C and D. The reasoning profile has the highest proportion of students and that relating to knowledge, the least proportion.

Furthermore, when profile performance bands are analysed across the four years, for the knowledge and comprehension profiles, a larger percentage of students obtained A and B grades than

C and D. For the knowledge profile, the average percentage of students who obtained A-B and C-D profile grades, respectively, was 79% and 21%, while for the comprehension dimension it was 55% and 45%. This difference was therefore, on average, substantially larger for the knowledge dimension (58%) than that of comprehension (10%). This suggests that for the knowledge profile, many more students achieved top band grades than for the comprehension profile.

Figure 5

Student Performance as Per Profiles for 2015, 2016, 2018, 2019

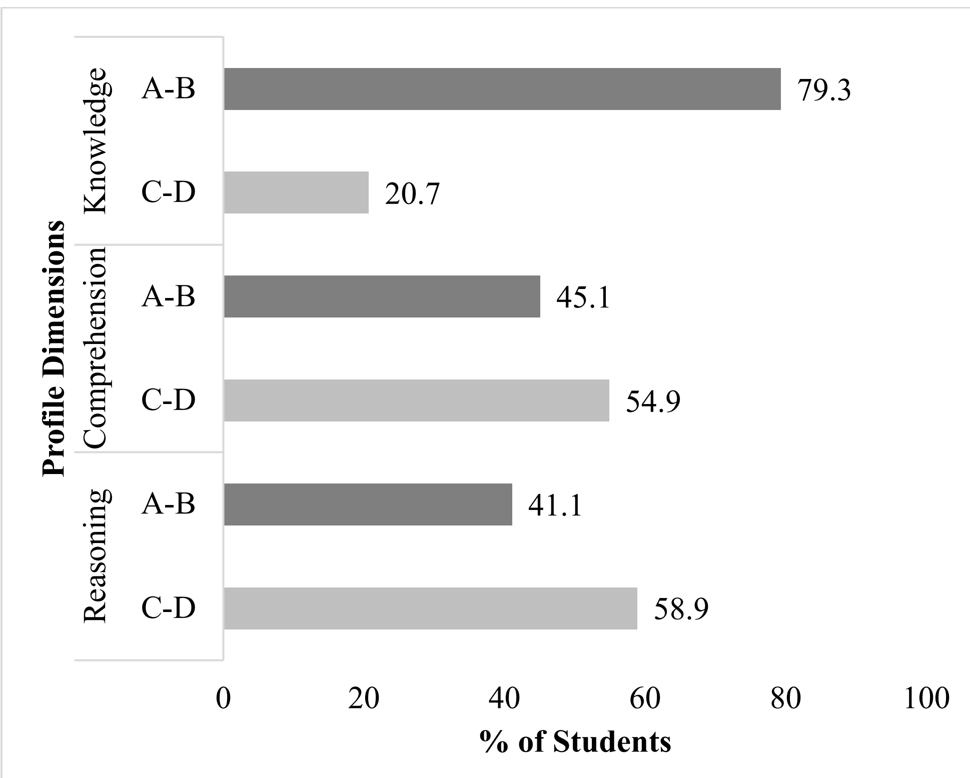


For the reasoning profile, this observed pattern in the knowledge and comprehension profiles was reversed with a larger percentage of students obtaining C and D profiles than A and B, for each of the four years. On average, across the four years, 59% of students obtained profiles within the satisfactory band (C-D) and 41% in the top band. For both the knowledge and comprehension dimensions, the percentage of students who obtained A-B and C-D profiles, respectively fluctuated from year to year. For the A-B band, there was a decrease from 2015-2016 (knowledge: -10; comprehension: -6) and 2018-2019 (knowledge: -26; comprehension: -5), and then an increase from 2016 to 2018 (knowledge: +19; comprehension: +3). For the reasoning profile, there was a small decrease in the percentage of students who obtained A-B profiles, year on year from 2015 to 2018 then a slight increase (+7) from 2018-2019. No student with a passing grade received E and F profile grades.

The pattern observed by year also holds for the entire data set. Figure 6 shows the percentage of students who achieved the profile bands A-B and C-D for each of the three profiles.

Figure 6

Student Performance as Per Profiles

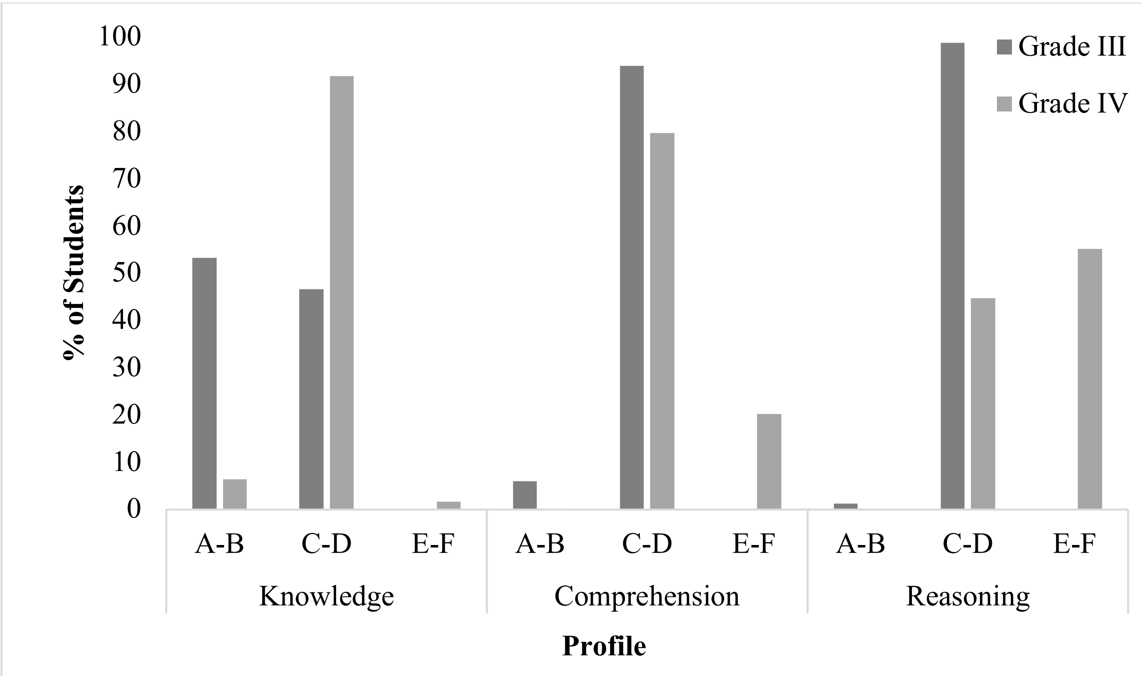


Research Question 3: How do the profile grade bands (knowledge, comprehension, reasoning) of students with a borderline failing grade (Grade IV) and a borderline passing grade (Grade III) compare?

Figure 7 shows the percentage of students with a borderline failing grade (Grade IV) and a borderline passing grade (Grade III) who attained different grade bands (A-B and C-D) for the knowledge, comprehension, and reasoning profiles.

Figure 7

Distribution of Students as per Profile Bands for Knowledge, Comprehension, and Reasoning



For the knowledge profile, the performance of students who obtain a Grade IV is substantially weaker than those who receive a borderline pass of Grade III. While for the former group, only 6.4% of students performed within the top band (A-B). For the latter, A-B were recorded for 53.3% of the students. The modal profile band for students attaining Grades III and IV, was A-B and C-D, respectively. Knowledge is the only profile for which students receiving a Grade IV had a top band (A-B) performance.

For the Comprehension profile, the C-D band, which is aligned to a satisfactory performance was the modal band of performance for both students obtaining a Grade III and IV. For students obtaining a Grade III, there was a large decrease in the percentage of students obtaining A and B profiles when compared to the Knowledge profile (52.4 versus 6.4). For these students, this trend extended to the reasoning profile, where only 1.2% of students achieved profile grades of A and B. For students who received a Grade IV, the percentage decrease (-6.3%) from the Knowledge profile to that of Comprehension was much smaller than that of the students who obtained Grade IIIs, perhaps because the percentage of students who obtained A and B profile grades for the former group was extremely small (6.4%). For the reasoning profile, no student who obtained a Grade IV had an A or B profile grade and the modal grade band was E-F. For students achieving a Grade III, for the reasoning profile almost all of the students (98.8%) achieved C and D profile grades, and no students obtained E and F grades. It appears that for both grade levels, students performed best and worst in the Knowledge and Reasoning profile, respectively.

Research Question 4: Is there a statistically significant difference in the profile grades of students with passing grades by gender?

Ho: There is no significant difference in the profile grades of males and females. An independent-samples *t*-test was conducted to compare the students' CSEC mathematics performance as per profiles by gender. For the knowledge and comprehension profiles, the mathematics performance of males was statistically significantly lower than females, however, for the reasoning profile the student gendered performance difference was not statistically significant. The gendered student performance for each profile is as follows: Knowledge - males (M = 5.1446, SD = 0.7526) and females (M = 5.1754, SD = 0.7567) $t(50243) = -4.495$, $p = .000$; Comprehension - males (M = 4.6542, SD = .89920) and females (M = 4.7252, SD = .90129) $t(50243) = -8.706$, $p = .000$; Reasoning - males (M = 4.3581, SD = .98962) and females (M = 4.3690, SD = 1.01929) $t(50243) = -1.191$, $p = .234$. The effect size related to the knowledge and comprehension profiles respectively were $d = 0.04$ and 0.08. According to Cohen et al. (2018), this effect size is tiny and a 4% and 8% difference was found between the two groups (males and females) for the knowledge and comprehension profiles respectively. This suggests that the difference between the groups, although statistically significant, was trivial.

Discussion

The data stemming from four years of CSEC mathematics examination administration and approximately 70,000 students saw that the students with the highest overall achievement also had the highest achievement across the profile domains. This is wholly in agreement with Mullis et al. (2012), who reported results related to the TIMMS mathematics component. This research also found a regular pattern of student profile performance for each of the grade levels (I-IV) and across the four years under consideration. Students performed best in the knowledge, then comprehension, and then reasoning profiles. Considering that a different cohort of students was assessed in each of the four years under study, this regularity in performance is interesting and noteworthy. This finding aligns with Mullis et al. (2012) who reported that for the 2011 administration of the TIMMS more countries demonstrated relative strengths in the knowing cognitive domain than the other two cognitive domains. It also agrees with Kaleli-Yılmaz and Hanci (2016) who found that 652 eighth grade Turkish students performed worst in the reasoning domain. Cato (2020) who investigated the CSEC mathematics results for a Caribbean sample also reported similar findings. However, results contrast with Kaleli-Yılmaz and Hanci (2016) who reported that their sample performed best in the applying domain. This finding from the present research also deviates from that of the 2019 offering of the TIMMS assessment for mathematics which indicated that students in more countries performed better in applying and reasoning than in knowing (Mullis et al., 2020). For an Australian cohort in the mathematics component of the TIMMS 2003, Thomson (2006) also found divergent results to the present study in that for both Years 4 and 8, comparatively, achievement in reasoning exceeded that of the other domains. For Year 4 and 8, respectively, achievement in applying and knowledge were lower compared to the other domains.

This study also found that the cognitive domains such as knowledge, comprehension, and reasoning have large, significant, positive relationships to mathematics achievement. This finding coincides with Pogoy et al. (2015), whose research related to the TIMMS 2011 data across countries. The author of this paper recognises that comparing CSEC student performance across multiple consecutive years with that of the TIMMS mathematics assessment that takes place every four years has its limitations. However, it is noted that in contrast to the consistent trend in performance across and within each of the four years that the analysis for this study revealed, for the TIMMS data, there appears to be a fluctuation in the cognitive domain that students in most countries perform best and worst in. For example, in 2011, candidate performance was superior in the knowing domain. This is in alignment with the present research whereas in 2019 it was the applying and reasoning domains. It may be that countries use the assessment results to improve student profile performance for the next TIMMS, hence the observed variation.

There are several explanations for the finding that students performed best in the knowledge, then comprehension, and reasoning profiles. The present author proposes several such hypotheses next. This finding may suggest, consistent with Mullis et al. (2003), that the performance in the knowledge profile dimension serves as the gatekeeper for students' ability to comprehend and reason. In the comprehension dimension, students apply procedures, concepts, and facts to routine problems. This dimension is akin to the applying dimension in the TIMMS. Mullis and Martin (2017) state that for students to engage optimally within this dimension, they need to possess knowledge of facts, concepts, and procedures, which they can then use in solving familiar mathematical problems involving different contexts. This indicates that knowledge may act as a prerequisite for application and reasoning (Pogoy et al., 2015; Thomson, 2006). This view is also supported by the National Council of Teachers of Mathematics (2014) who states that learners should "acquire conceptual knowledge as well as procedural knowledge, so that they can meaningfully organize their knowledge …and transfer and apply knowledge to new situations" (p. 9). The current finding and these explanations could lend support to the argument that was briefly introduced earlier regarding whether Bloom's taxonomy was hierarchical or not. They appear to support the notion that the taxonomy may be hierarchical, but future research is needed to further delineate this. This contrasts with Kilpatrick et al. (2001) who appear to prioritise reasoning over knowledge and comprehension although acknowledging that all cognitive dimensions are important and mutually influential. This research provides empirical evidence to support the previously mentioned assertion of Kilpatrick et al. (2001) and is significant in that it adds to the existing literature, new ways in which the profiles or cognitive domains are related to each other. Reasoning has also been found to help students see connections between their present and prior knowledge thereby augmenting the prospect of retaining and recalling this knowledge (Ball & Bass, 2003; National Council of Teachers of Mathematics, 2009). Consequently, it is plausible that the group of students in this research who are competently able to reason are using this to boost their knowledge and comprehension scores.

Other explanations related to the research context may explain this observed pattern that the students who participated in this research performed best in the knowledge, then comprehension, and reasoning profiles. In Jamaica, like international jurisdictions, both policy and curricular documents emphasise that mathematics teaching should aim to develop analytical, reasoning, and critical thinking skills (Jeannotte & Kieran, 2017; Ministry of Education, 2013, 2017). However, the yearly examination reports, based on the performance of test-takers in the CSEC mathematics examinations, consistently note that in general, candidates' reasoning and problem-solving skills were wanting and made recommendations relating to instruction to teachers in that regard (Caribbean Examination Council, 2020). Further to this, in Jamaica, the current curriculum in use in Grades 7-9 was recently developed based on the finding in early reform reports that the curriculum and related teaching were too focused on the retention of factual knowledge and not on the development of transferable skills and competences (Ministry of Education, 2016). Consequently, it is not wholly surprising that the reasoning cognitive domain would be the weakest since it has been given the least emphasis in practice. Stigler and Hiebert (2009) describe teaching as a cultural activity which "evolve(s) over long periods of time in ways that are consistent with the stable web of beliefs and assumptions that are part of the culture" (p. 87). Furthermore, they point out that the instructional approaches that most teachers apply to their practice, closely resemble those that were used by their own teachers when they were students and in earlier times as well. Based on the aforementioned, it is reasonable to conclude that the teaching methods that present teachers tasked with developing students' reasoning and problem-solving skills were exposed to would have seldom focused on the development of reasoning or critical thinking skills. Instead, they would very likely have mainly centred on the acquisition of knowledge and skills. This would suggest that it is plausible that many present-day Jamaican teachers would have weak reasoning skills in mathematics and so would likely find it difficult to facilitate their development in mathematics lessons. Previous research (e.g., Mata-Pereira & Da Ponte, 2017; Richland et al., 2012) has linked the development of students' reasoning capabilities directly to the mathematics instruction that they receive. In this regard, Brodie (2009) states that "mathematics reasoning is challenging to learn and teach. For teachers who learned mathematics and learned to teach mathematics in traditional ways, the challenges are enormous" (p. 3). Based on this contextual consideration, the finding that the reasoning cognitive domain would be the weakest is reasonable.

These new findings are wholly plausible and in line with Camilli and Dossey (2019) who note that "educational systems of different participants have distinct profiles of mathematics achievement" (p. 1). Consequently, the present author acknowledges that the findings related to student profile performance, and the relationship between achievement grades and that related to profile/cognitive domains, presented here do not constitute a typical picture of high school mathematics achievement in general. For a different sample of high school students, other patterns of performance could emerge. The findings are noteworthy since this study reveals a regular pattern of profile performance across achievement grade levels and multiple years of the administration of the high stakes, large-scale mathematics examination. This has not been reported in previous research or large-scale assessments in mathematics (e.g., Kaleli-Yılmaz & Hanci, 2016; Mullis et al., 2020). In this regard, the findings from this study add to the mixed findings related to the student mathematics performance as per cognitive domains. The present findings also provide new evidence relating to the potential impact of the knowledge domain on the other two profiles/cognitive dimensions and the relationship between the profiles and mathematics achievement.

The disaggregation by years of the profile grade band performance of students who obtained passing grades revealed that for each of the years being considered in this research (2015, 2016, 2018, 2019), for the knowledge and comprehension profiles, a larger percentage of students obtained A and B grades than C and D. This pattern is reversed for the reasoning profile. This is a cause for concern for stakeholders in education in Jamaica since the island, for the first time, will participate in the next administration of the Programme for International Student Assessment (PISA) in 2022 (Hunter, 2019). This large-scale, international assessment measures 15-year-old students' ability to apply their knowledge and skills related to reading, mathematics, and science in real world problem solving situations (*Programme for International Student Assessment*, 2021). A major focus of the mathematics component of this assessment is mathematical reasoning due to the central role it plays in mathematical literacy in the 21st century (OECD, 2019). While the CSEC mathematics examination data involves 16-year-old students and the PISA 15-year-olds, the CSEC examination data suggests that Jamaican students may struggle with the mathematics component of the PISA. The aforementioned presents an example as to how a country's national data can be used to provide insights that could improve educational programs, establishments, or systems, enhance practices in education, and support the learning processes of individuals (Koeppen et al., 2008) and ultimately help prepare students for an international assessment. This approach if not yet utilized can be replicated in other jurisdictions.

In the 2016-17 academic year, a new national curriculum that emphasises the development of higher order thinking skills such as reasoning in mathematics, was implemented in Jamaican primary and lower secondary schools. The previously mentioned findings could form the basis on which to monitor changes in achievement in the cognitive domains, post implementation of the new curriculum. Changes in achievement may reflect the changes made in the national curriculum in Jamaica. This confirms Mullis et al. (2012) and serves as yet another example of the key role that assessment can play within educational reform. Future research could explore achievement in the cognitive domains post this implementation.

This research found that there is a significant gendered performance gap in favour of the females as it relates to the three profile dimensions of knowledge, comprehension, and reasoning. This finding coincides with previous work related to CSEC mathematics achievement in Jamaica (George, 2022) and international research relating to mathematics assessments (e.g., Cimpian et al., 2016;

Lindberg et al., 2010; Lubienski & Pinheiro, 2020). Gender disparity in education is unfavourable from the perspectives of educational or social justice and equity. For this reason, among others, the Sustainable Development Goals, an international plan of action proposed by The United Nations and adopted by many countries including Jamaica, as a global agenda aims to "eliminate gender disparities in education" (UNESCO, 2021). As a result, there is a research imperative to uncover the presence of these disparities and to decrease them, with the aim of ultimately eradicating their occurrence (Evans, 1997; OECD, 2015; UNESCO, 2017). In this regard, the findings of this study related to gender differences favouring females are important. Mathematics plays a crucial role in an individual's scholastic and vocational success. Also, it is an essential component in fields related to Science, Technology, Engineering and Mathematics (STEM) which are instrumental in advancing innovation and competition in the 21st century (Marginson et al., 2013; Waite & McDonald, 2019). It also promotes financial productivity and development for a country's citizens, industries, and businesses (Ukpata & Nancy, 2012). Evans (1997) notes that gender differences in education highlight human capital issues since developing countries need educated persons with a wide range of knowledge, skills, and competencies.

Although this research found significant gender differences in mathematics achievement the related effect sizes were small and so from a practical perspective this difference is negligible. This difference is also not as remarkable as has been found in other Caribbean territories, such as Trinidad & Tobago (OECD, 2016) who reported the largest gender gap in favour of females in the 72 countries who participated in the 2015 PISA mathematics component. Notwithstanding the small effect size, Stoet et al. (2016) asserted that the gender disparity is important and "one of the main psychological and educational research aims is to determine which factors can explain the sex difference in mathematics performance" (p. 2). Future research could seek to explore this issue in greater depth. The present finding also contrasts with Stoet and Geary (2013) who analysed data from four administrations of PISA and found that in most countries, boys scored higher than girls in mathematics assessments.

Conclusion

Mullis et al. (2000) asserted that "it is important that educators, curriculum developers, and policy makers understand what students know and can do in mathematics and what areas, … need more focus and effort" (pg. 57). In this regard, this study explored different facets of students' mathematics achievement as per three cognitive domains, *knowledge, comprehension, and reasoning,* to provide educational stakeholders with a more fulsome and deeper understanding of this achievement. In particular, it investigated student profile performance in the CSEC mathematics examinations as per overall achievement grades (I-IV), across four years (2015, 2016, 2018, 2019), and by gender. It also explored the relationship between the profiles and mathematics achievement.

This study is significant in several ways. First, it found a new and regular pattern of cognitive domain performance across multiple years for different cohorts of students. Students performed best in the knowledge, then comprehension, and reasoning profiles. The Jamaican mathematics policy and curricular, in alignment with current research and international practices within mathematics education, emphasise that the development of reasoning skills is a central focus of instruction. This research shows that this goal is not currently being attained. Students appear to have a relative weakness in reasoning which signifies to teachers and other educational stakeholders, such as teacher educators, an area that requires more effort and focus as per Mullis et al. (2000).

Also, students with a borderline failing grade had notable deficiencies in the knowledge domain and that the homogeneity of student profile performance decreased with each grade level from I-IV. This indicates that perhaps a wider variety of resources and/or pedagogical approaches including the use of differentiated instruction may be needed for engaging with lower performing students than for high performing students. Teachers could aim to improve students' performance in the comprehension/applying domains by investing instructional time in improving students' knowledge of mathematical facts, concepts, and procedures. This is particularly important for low performing students whose low performance in the knowledge domain appears to impact their performance in the other two domains.

This research also found that there is a strong, positive relationship between achievement as per the three cognitive domains and general mathematics achievement. Therefore, a focus on strengthening student mathematics achievement in the three cognitive dimensions is likely to result in improved student achievement. For Jamaica, this would be a key educational imperative considering the longstanding poor achievement of students in mathematics. This could be prioritised especially in the earlier grades of high school so that adequate time can be spent on facilitating the development in all the cognitive domains. In order to implement the aforementioned recommendations relating to teaching, there is the need for training of teachers in developing students' higher order thinking skills. This could be undertaken by teachers' training institutions or the central Ministry of Education.

This research focused on the cognitive domains related to students' mathematics achievement. Since performance outcomes can be classified with regard to content and cognitive domains, future research could centre on student achievement, as per the content domains assessed in the CSEC mathematics examination. That research could be undertaken using the raw scores of students to overcome a limitation of the current research, that is, to analyse overall grades. Since teachers are expected to facilitate the development of students' competencies regarding the cognitive domains, future research could also explore mathematics teachers' existing knowledge and skills relating to the three cognitive domains to ascertain whether and where gaps exist that need to be addressed.

The author received no financial support for the research, authorship, and/or publication of this manuscript.

Lois George (lois.george@open.uwi.edu) is a Senior Lecturer in Mathematics Education in the School of Education at The University of the West Indies (UWI), Mona Campus. She has graduate degrees in Mathematics Education from the University of Southampton and Measurement, Testing, and Evaluation from UWI. Her research interests include mathematical cognitive development, the affective domain in mathematics education, and different elements of mathematics assessments.

References

Amrein-Beardsley, A., & Berliner, D. C. (2002). *An analysis of some unintended and negative consequences of high-stakes testing*. Education Policy Research Unit. https://nepc.colorado.edu/sites/default/files/EPSL-0211-125-EPRU.pdf

Ball, D. L., & Bass, H. (2003). Making mathematics reasonable in school. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 27-44). National Council of Teachers of Mathematics. https://www.researchgate.net/publication/312532588_Making_mathematics_reasonable_in_school

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1 - Cognitive Domain*. David McKay Company, Inc.

Bradbury, A., Braun, A., & Quick, L. (2021). Intervention culture, grouping and triage: high-stakes tests and practices of division in English primary schools. *British Journal of Sociology of Education, 42*(2), 147-163. https://doi.org/10.1080/01425692.2021.1878873

Brodie, K. (2009). *Teaching mathematical reasoning in secondary school classrooms* (Vol. 775). Springer Science & Business Media.

Camilli, G., & Dossey, J. A. (2019). Multidimensional national profiles for TIMSS 2007 and 2011 mathematics. *The Journal of Mathematical Behavior, 55*, 100693. https://doi.org/10.1016/j.jmathb.2019.02.001

Caribbean Examination Council. (2020). *CSEC mathematics syllabus, specimen paper, mark scheme, subject reports*. Caribbean Examination Council (CXC).

Caribbean Examinations Council. (2015). *Caribbean secondary education certificate (CSEC) mathematics syllabus effective for examinations from May–June 2018*. Caribbean Examinations Council (CXC). https://www.cxc.org/SiteAssets/CSEC_Mathematics_Syllabus_with_Specimen_Papers.pdf.

Carifio, J., & Perla, R. (2008). Resolving the 50-year debate around using and misusing Likert scales. *Medical Education, 42*(12), 1150-1152. https://doi.org/10.1111/j.1365-2923.2008.03172.x

Cato, B. R. (2020). *Mathematics cognitive and content abilities across the Vincentian student population* [Doctoral dissertation, Walden University]. ScholarWorks. https://scholarworks.waldenu.edu/cgi/viewcontent.cgi?article=10751&context=dissertations

Cimpian, J. R., Lubienski, S. T., Timmer, J. D., Makowski, M. B., & Miller, E. K. (2016). Have gender gaps in math closed? Achievement, teacher perceptions, and learning behaviors across two ECLS-K cohorts. *AERA Open, 2*(4), 1-19. https://doi.org/0.1177/2332858416673617

Crossfield, D., & Bourne, P. A. (2017). Education professionals' perceptions of factors that contribute to effective mathematics teaching and achievement in Jamaica. *International Journal of Research in Humanities and Social Studies, 4*(12), 1-10. https://www.ijrhss.org/papers/v4-i12/1.pdf

Department for Education. (2013). *Mathematics: GCSE subject content and assessment objectives*. Department for Education. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/254441/GCSE_mathematics_subject_content_and_assessment_objectives.pdf

Dogan, E., & Tatsuoka, K. (2008). An international comparison using a diagnostic testing model: Turkish students' profile of mathematical skills on TIMSS-R. *Educational Studies in Mathematics, 68*(3), 263-272. https://doi.org/10.1007/s10649-007-9099-8

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136*(1), 103-127. https://doi.org/10.1037/a0018053

Evans, H. (1997). *Gender differences in education in Jamaica.* World Council of Comparative and International Educational Societies.

Fenanlampir, A., Batlolona, J. R., & Imelda, I. (2019). The struggle of Indonesian students in the context of TIMSS and PISA has not ended. *International Journal of Civil Engineering and Technology, 10*(2), 393-406. https://paper.researchbib.com/view/paper/210729

Forgasz, H. J. (2012, July). Gender and mathematics in Australia: A downward trajectory. In S. J. Cho (Ed.), Proceedings of the 12th International Congress on Mathematical Education, Seoul, Korea. https://link.springer.com/content/pdf/10.1007/978-3-319-12688-3.pdf

George, L. (2020). Exploring the m in stem: Post-secondary participation, performance and attrition in mathematics. *Canadian Journal of Science, Mathematics, and Technology Education, 20*, 3, 441–461. https://doi.org/10.1007/s42330-020-00095-6

George, L. (2022). Investigating the link between poor mathematics achievement and youth marginalization in Jamaica. In S. Blackman (Ed.), *Equitable education for marginalized youth in the caribbean and Latin America* (pp. 105-127). Routledge.

George, P. (2013). Made for mathematics? Implications for teaching and learning. In D. Leslie & H. Mendick (Eds.), *Debates in mathematics education* (pp. 58-66). Routledge. https://doi.org/10.4324/9780203762585

Griffith, S. A. (2013). Choice and performance in CSEC and CAPE TVET subjects: A comparison with more conventional subjects. *Caribbean Curriculum, 21*, 97-119. https://uwispace.sta.uwi.edu/server/api/core/bitstreams/61e100a9-2f3a-40bf-b8f8-2a34771254b4/content

Harks, B., Klieme, E., Hartig, J., & Leiss, D. (2014). Separating cognitive and content domains in mathematical competence. *Educational assessment, 19*(4), 243-266. https://doi.org/10.1080/10627197.2014.964114

Hefty, L. J. (2015). STEM gives meaning to mathematics. *Teaching Children Mathematics, 21*(7), 422-429. https://doi.org/10.5951/teacchilmath.21.7.0422

Hunter, J. A. (2019, July). *Students to take part in PISA in 2021*. Jamaica Information Service. https://jis.gov.jm/students-to-take-part-in-pisa-in-2021/

Hutchings, M. (2015). *Exam factories? The impact of accountability measures on children and young people*. https://www.basw.co.uk/system/files/resources/basw_112157-4_0.pdf

Jeannotte, D., & Kieran, C. (2017). A conceptual model of mathematical reasoning for school mathematics. *Educational Studies in Mathematics, 96*(1), 1-16. https://doi.org/10.1007/s10649-017-9761-8

Kaleli-Yılmaz, G., & Hanci, A. (2016). Examination of the 8th grade students' TIMSS mathematics success in terms of different variables. *International Journal of Mathematical Education in Science and Technology, 47*(5), 674-695. https://doi.org/10.1080/0020739X.2015.1102977

Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding+ it up: Helping children learn mathematics*. National Academies Press. https://www.nap.edu/download/9822

Koeppen, K., Hartig, J., Klieme, E., & Leutner, D. (2008). Current issues in competence modeling and assessment. *Zeitschrift für Psychologie/Journal of Psychology, 216*(2), 61-73. https://doi.org/10.1027/0044-3409.216.2.61

Leder, G. C. (2012, July). Gender and mathematics education revisited. In S. J. Cho (Ed.), Proceedings of the 12th International Congress on Mathematical Education, Seoul, Korea. https://link.springer.com/content/pdf/10.1007/978-3-319-12688-3.pdf

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin, 136*(6), 1123–1135. https://doi.org/10.1037/a0021276

Lubienski, S., & Pinheiro, W. A. (2020). Gender and mathematics: What can other disciplines tell us? What is our role? *Journal of Urban Mathematics Education, 13*(1), 1–14. https://files.eric.ed.gov/fulltext/EJ1254528.pdf

Marginson, S., Tytler, R., Freeman, B., & Roberts, K. (2013). *STEM: Country comparisons: International comparisons of science, technology, engineering and mathematics (STEM) education. (Final Report No. 0987579800)*. Australian Council of Learned Academies. https://acola.org/wp-content/uploads/2018/12/saf02-stem-country-comparisons.pdf

Marks, R. (2014). Educational triage and ability-grouping in primary mathematics: A case-study of the impacts on low-attaining pupils. *Research in Mathematics Education, 16*(1), 38-53. https://doi.org/10.1080/14794802.2013.874095

Mata-Pereira, J., & Da Ponte, J. P. (2017). Enhancing students' mathematical reasoning in the classroom: teacher actions facilitating generalization and justification. *Educational Studies in Mathematics, 96*(2), 169-186. https://doi.org/10.1007/s10649-017-9773-4

McMahon, R. (2022). *An examination of Florida's high stakes testing: a post-hoc study of the academic achievement of students with borderline intellectual functioning*. [Doctoral dissertation, University of Central Florida]. Showcase of Text, Archives, Research, and Scholarship. https://stars.library.ucf.edu/cgi/viewcontent.cgi?article=2255&context=etd2020

McNeil, L. (2002). *Contradictions of school reform: Educational costs of standardized testing*. Routledge.

McPherson, A. (2020, November 13). *Understanding your grades [Video]. YouTube.* https://www.youtube.com/watch?v=zUKBglpJW3s

Minarechová, M. (2012). Negative impacts of high-stakes testing. *Journal of Pedagogy, 3*(1), 82-100. https://doi.org/10.2478/v10159-012-0004-x

Ministry of Education, Youth and Information [Jamaica]. (2013). *National mathematics policy guidelines.* https://moey.gov.jm/wp-content/uploads/2015/05/National-Mathematics-Policy-Guidelines-2013.pdf

Ministry of Education, Youth and Information [Jamaica]. (2016). *Ministry paper #47/16: The national standards curriculum (NSC).* https://www.japarliament.gov.jm/attachments/article/1678/2016%20Ministry%20Paper%2047.pdf

Ministry of Education, Youth and Information [Jamaica]. (2017). *National standards curriculum grades 7-9 mathematics*. Ministry of Education, Youth and Information.

Mullis, I. V., & Martin, M. O. (Eds.). (2017). *TIMSS 2019 assessment frameworks*. TIMSS & PIRLS International Study Center, Boston College. http://timssandpirls.bc.edu/timss2019/frameworks/.

Mullis, I. V., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics*. TIMSS & PIRLS International Study Center.

Mullis, I. V., Martin, M. O., Smith, T. A., Garden, R. A., Gregory, K. D., Gonzalez, E. J., Chrostowski, S. J., & O'Connor, K. M. (2003). *TIMSS Assessment frameworks and specifications 2003, 2nd Edition*. International Association for the Evaluation of Educational Achievement. https://timss.bc.edu/timss2003i/PDF/t03_af_book.pdf

Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2016). *TIMSS 2015 international results in mathematics*. TIMSS & PIRLS International Study Center. http://timssandpirls.bc.edu/timss2015/international-results/wp-content/uploads/filebase/full%20pdfs/T15-International-Results-in-Mathematics.pdf.

Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019 international results in mathematics and science*. TIMSS & PIRLS International Study Center, Boston College. https://timss2019.org/reports/wp-content/themes/timssandpirls/download-center/TIMSS-2019-International-Results-in-Mathematics-and-Science.pdf

National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all.* National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (2009). *Executive summary -Focus in high school mathematics: Reasoning and sense making*. National Council of Teachers of Mathematics. https://www.nctm.org/uploadedFiles/Research_and_Advocacy/Advocacy_Toolkit/FHSM%20Executive%20Summay.pdf

Nelson, G., & Powell, S. R. (2018). A systematic review of longitudinal studies of mathematics difficulty. *Journal of Learning Disabilities, 51*(6), 523-539. https://doi.org/10.1177/0022219417714773

Nilsen, T., Gustafsson, J.-E., & Blömeke, S. (2016). *Conceptual framework and methodology of this report* (Vol. 2). International Association for the Evaluation of Educational Achievement (IEA). https://link.springer.com/chapter/10.1007/978-3-319-41252-8_1

OECD. (2015). *The ABC of gender equality in education: Aptitude, behaviour, confidence.* https://bit.ly/31KZyMO

OECD. (2016). *PISA 2015 results excellence and equity in education (Volume I).* https://publications.iadb.org/publications/english/document/Latin-America-and-the-Caribbean-in-PISA-2015-How-Do-Boys-and-Girls-Perform.pdf

OECD. (2019). *PISA 2018 Assessment and Analytical Framework*. https://doi.org/10.1787/b25efab8-en

Pogoy, A., Balo, V. T., Obaob Jr, G., & Chiu, S. (2015). Fractal correlations on content and cognitive domains and mathematics performance across countries. *European Scientific Journal, 11*(16), 344-352. https://core.ac.uk/download/pdf/236405877.pdf

Postlethwaite, T. N. (1994). Validity vs. utility: Personal experiences with the taxonomy. In L. W. Anderson & L. A. Sosniak (Eds.), *Bloom's taxonomy: A forty year retrospective, Ninety third yearbook of the national society for the study of education, part II* (pp.174-180). The National Society for the Study of Education.

Programme for International Student Assessment. (2021). OECD. https://www.oecd.org/pisa/

Reback, R. (2008). Teaching to the rating: School accountability and the distribution of student achievement. *Journal of Public Economics, 92*(5-6), 1394-1415. https://doi.org/10.1016/j.jpubeco.2007.05.003

Richland, L. E., Stigler, J. W., & Holyoak, K. J. (2012). Teaching the conceptual structure of mathematics. *Educational Psychologist, 47*(3), 189-203. https://doi.org/10.1080/00461520.2012.667065

Rothman, T., & Henderson, M. (2011). Do school-based tutoring programs significantly improve student performance on standardized tests? *RMLE Online, 34*(6), 1-10. https://doi.org/10.1080/19404476.2011.11462079

Schoenfeld, A. H., & Kilpatrick, J. (2008). Toward a theory of proficiency in teaching mathematics. In S. Llinares & O. Chapman (Eds.), *International handbook of mathematics teacher education*, *Volume 2* (pp. 321-354). Brill Sense.

Shalem, Y., Sapire, I., & Huntley, B. (2013). Mapping onto the mathematics curriculum-an opportunity for teachers to learn. *Pythagoras, 34*(1), 1-10. https://doi.org/10.4102/pythagoras.

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M. I., & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science, 23*(7), 691-697. https://doi.org/10.1177/0956797612440101.

Spencer-Ernandez, J., & George. (2016). Single sex vs. co-educational high schools: Performance of Caribbean students across school types in mathematics on the Caribbean Secondary Education Certificate. *Caribbean Education Research Journal, 4*(2), 96-121. https://bit.ly/2WrvNOY

Steiner, E. T., & Ashcraft, M. H. (2012). Three brief assessments of math achievement. *Behavior Research Methods, 44*(4), 1101-1107. https://doi.org/10.3758/s13428-011-0185-6

Stigler, J. W., & Hiebert, J. (2009). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. Free Press.

Stoet, G., Bailey, D. H., Moore, A. M., & Geary, D. C. (2016). Countries with higher levels of gender equality show larger national sex differences in mathematics anxiety and relatively lower parental mathematics valuation for girls. *PloS one, 11*(4). https://doi.org/10.1371/journal.pone.0153857

Stoet, G., & Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: Within-and across-nation assessment of 10 years of PISA data. *PloS one, 8*(3), e57988. https://doi.org/10.1371/journal.pone.0057988

Suurtamm, C., Thompson, D. R., Kim, R. Y., Moreno, L. D., Sayac, N., Schukajlow, S., Silver, E., Ufer, S., & Vos, P. (2016). *Assessment in mathematics education: Large-scale assessment and classroom assessment*. Springer.

Thomson, S. (2006). Australian students achievement in the TIMSS 2002 mathematics cognitive domains. *TIMSS Australia Monograph Series*. https://research.acer.edu.au/cgi/viewcontent.cgi?article=1001&context=timss_monographs

Ukpata, S. I., & Nancy, A. (2012). Mathematics as a tool in human capital formation and development in Nigeria. *Journal of Economics, 3*(2), 95-107. https://doi.org/10.1080/09765239.2012.11884956

UNESCO. (2017). *A guide for ensuring inclusion and equity in education*. https://unesdoc.unesco.org/ark:/48223/pf0000248254

UNESCO. (2021). *Sustainable Development Goal 4 and its targets*. https://en.unesco.org/education2030-sdg4/targets

Waite, A. M., & McDonald, K. S. (2019). Exploring challenges and solutions facing STEM careers in the 21st century: A human resource development perspective. *Advances in Developing Human Resources, 21*(1), 3-15. https://doi.org/10.1177/1523422318814482

Webb, D. C. (2012, July). Teacher change in classroom assessment: The role of teacher content knowledge in the design and use of productive classroom assessment. In S. J. Cho (Ed.), Proceedings of the 12th International Congress on Mathematical Education: Topic Study Group 33, Seoul, Korea.

Webb, D. C. (2020). *Bloom's taxonomy in mathematics education* (2nd ed.). Springer.

Wu, H., & Leung, S. O. (2017). Can Likert scales be treated as interval scales?—A simulation study. *Journal of Social Service Research, 43*(4), 527-532. https://doi.org/10.1080/01488376.2017.1329775

Appendix A

Figure 1

Specimen Caribbean Secondary Education Certificate Mathematics Question 2 (Caribbean Examinations Council, 2015, p. 87)

(a) Simplify: $p^3q^2 \times pq^5$ **(1 mark)**

(b) If $a * b = 2a - 5b$, calculate the value of

(i) $3 * 4$ **(1 mark)**

(ii) $(3 * 4) * 1$ **(1 mark)**

(c) Factorize completely: $3x + 6y - x^2 - 2xy$ **(2 marks)**

(d) A string of length 14 cm is cut into two pieces. The length of the first piece is x cm. The second piece is 5 cm **longer** than half the length of the first piece.

(i) State in terms of x, the length of the second piece of string. **(1 mark)**

(ii) Write an expression, in terms of x, to represent the TOTAL length of the two pieces of string. **(1 mark)**

(iii) Hence, calculate the length of the first piece of string. **(2 marks)**

Total 9 marks

Appendix B

Figure 2

Profiles Assigned to the Solutions to Question 2 (Caribbean Examinations Council, 2015, p. 99)

	PROFILES K	PROFILES C	PROFILES R	Total
(a) $p^3q^2 \times pq^5 = p^4q^7$	1			
	1	**–**	**–**	**1**
(b) $a * b = 2a - 5b$				
(i) $3 * 4 = 2(3) - 5(4)$ $= 6 - 20$ $= -14$		1		
(ii) $(3 * 4) * 1 = (-14) * 1$ $= 2(-14) - 5(1)$ $= -28 - 5$			1	
$= -33$	**–**	**1**	**1**	**2**
(c) $3x + 6y - x^2 - 2xy$				
$= 3(x + 2y) - x(x + 2y)$		1		
$= (3 - x)(x + 2y)$	1			
	1	**1**	**–**	**2**
(d) (i) Length of 2nd piece $= \frac{1}{2}x + 5$		**1**		
(ii) Sum of two lengths $= x + \frac{1}{2}x + 5$ $= \frac{3}{2}x + 5$	**1**			
(iii) $= \frac{3}{2}x + 5 = 14$ $= \frac{3}{2}x = 9$ $x = 6$			**1**	
Length of 1st piece is 6 cm		**1**		
	1	**2**	**1**	**4**
TOTAL	**3**	**4**	**2**	**9**

Appendix C

Figure 3

Specimen Caribbean Secondary Education Certificate Mathematics Question 4 (Caribbean Examinations Council, 2015, p. 89)

(a) The line BC passes through the point A(−5, 3) and has a gradient of $\frac{2}{5}$.

(i) Write the equation of the line BC in the form $y = mx + c$. **(2 marks)**

(ii) Determine the equation of the line which passes through the origin and is perpendicular to the line BC. **(2 marks)**

(b) The functions f and g are defined as:

$f(x) = \frac{2x - 1}{x + 3}$ $g(x) = 4x - 5$

(i) Determine f g(3). **(2 marks)**

(ii) Derive an expression for $f^{-1}(x)$. **(3 marks)**

Total 9 marks

Appendix D

Figure 4

Profiles Assigned to the Solutions to Question 4 (Caribbean Examinations Council, 2015, p. 103)

	PROFILES			Total
	K	C	R	
(a) A(−5, 3); $m = \frac{2}{5}$				
(i) Equation of line BC is $y - 3 = \frac{2}{5}(x + 5)$		1		
i.e. $y = \frac{2}{5}x + 5$		1		
(ii) Gradient of line perpendicular to BC is $-\frac{5}{2}$	1			
Equation of line through (0, 0) is $y = -\frac{5}{2}x$			1	
	1	2	1	4
(b) $f(x) = \frac{2x-1}{x+3}$; $g(x) = 4x - 5$				
(i) $g(3) = 12 - 5 = 7$	1			
$\therefore fg(3) = \frac{14 - 1}{7 + 3} = \frac{13}{10}$			1	
(ii) $\frac{2(f^{-1})-1}{(f^{-1})+3} = x$			1	
$2(f^{-1}) - 1 = xf^{-1} + 3x$		1		
$f^{-1}(2 - x) = 3x + 1$				
$f^{-1} = \frac{3x + 1}{2 - x}$	1			
	2	1	2	5
TOTAL	3	3	3	9

ELECTRONIC JOURNAL FOR RESEARCH
IN SCIENCE & MATHEMATICS EDUCATION
VOL. 27, NO. 3, 64-74

The Teacher is Key to STEM Education for All: A Catalyst for Competitive Workforce and Economic Development

David Devraj Kumar
Florida Atlantic University

Susannah L. Brown
Florida Atlantic University

ABSTRACT

The teacher is key to reforming K-12 science, technology, engineering, and mathematics (STEM) education for all students (Powell et al., 2018) in the United States, and a catalyst for a competitive workforce and economic development. Reports based on Free or Reduced-Price Lunch (FRPL) and percent schools offering STEM courses, along with FRPL and per pupil expenditure in science adjusted for inflation, show disparity between the highest and lowest quartiles (U.S. Government Accountability Office, 2018; Banilower et al., 2018). Strategies to promote STEM for All and turn STEM education into a dependable human resource pipeline for a competitive workforce and economic development are discussed. The strategies include, promoting diversity and inclusion towards STEM for All, providing adequate STEM teacher training, increasing teacher retention in STEM subjects, and building a supportive environment for STEM teachers and teachers in general. These strategies are essential to connecting K-12 STEM education, a competitive workforce, and economic development.

Keywords: K-12 education, STEM education, competitive workforce, economic development, FRPL, inclusion, diversity, teacher training, teacher retention

Introduction

Strategies for reforming Kindergarten through grade twelve (K-12) science, technology, engineering, and mathematics (STEM) education in the United States as a dependable human resource pipeline for a competitive workforce and economic development are explored. Selected STEM and STEM education information and data from the U.S. contribute to this discussion. Connecting STEM education, workforce readiness, and the economy is a trend not only in the U.S. but also a global trend (World Economic Forum, 2020). According to the U.S. Department of Commerce, Economics and Statistics Administration, Office of the Chief Economist (OCE) (Noonan, 2017), STEM workers play a significant part in "innovation and competitiveness by generating new ideas and new companies" (p. 1).

Accordingly, the OCE (Noonan, 2017) reports that in the private sector, the average hourly wage of a STEM worker based upon educational level is higher than a non-STEM worker as presented in Table 1. For example, STEM hourly wage average ($27.53) for workers with a high school diploma or less is 69.8%, over a similar non-STEM worker average hourly wage ($16.21) (Noonan, 2017). Additionally, for STEM workers with less than a graduate degree, regression-based hourly earnings premiums have increased over time since the mid-1990s compared to non-STEM workers with graduate degrees (Noonan, 2017). Calculations by OCE (Noonan, 2017) based on Bureau of Labor Statistics Employment Projections and the National Bureau of Economic Research, indicate the projected growth in STEM employment is 8.9% compared to non-STEM jobs during 2014-2024.

Table 1

Private Sector Average Hourly Wage (USD) by Educational Level, STEM vs. non-STEM

Educational Level	STEM Hourly Wage Average (USD)	Non-STEM Hourly Wage Average (USD)	Percentage Increase for STEM Hourly Wage Average (USD)
High School Diploma or less	$27.53	$16.21	69.8%
Associate Degree or Some College	$30.79	$19.09	61.3%
Bachelor's Degree	$39.28	$28.34	38.6%
Graduate Degree	$45.37	$35.16	29.0%

Note. Office of Chief Economist calculations based on Current Population Survey public-use data. (Noonan, 2017)

According to the U.S. Department of Labor, Bureau of Labor Statistics (2017), in May 2015, 8.6 million (6.2%) U.S. jobs were in STEM fields, of which 750,000 were in applications software development and 333,010 were in wholesale and manufacturing sales of scientific and technical products. When looking at these statistics, it is not a surprise that the Committee on STEM Education of the National Science and Technology Council (2018) outlined the following three goals in the document entitled *Charting a Course for Success: America's Strategy for STEM Education*: 1) Build solid foundations for STEM literacy, 2) Prepare the STEM workforce for the future, and 3) Increase work-based learning and training through educator-employer partnerships. As usual, along with a plethora of similar ambitious rhetoric, the Committee on STEM Education of the National Science and Technology Council (2018) report failed to address significant strategies key to strengthening STEM education at the K-12 level. To maintain a steady pipeline of STEM-trained human resources, it is necessary to promote STEM subjects in K-12 schools. Facilitating STEM education early at the elementary school level is an especially important strategy supporting the first goal of building solid foundations for STEM literacy (Committee on STEM Education of the National Science and Technology Council, 2018; Oberoi, 2016). To further support this goal, strategies focusing on K-12 STEM education such as, promoting diversity and inclusion towards STEM for All; providing adequate STEM teacher training; increasing teacher retention in STEM subjects; and building a supportive environment for STEM teachers as well as teachers in general, are discussed.

What is STEM Education?

STEM education is an integrated approach to teaching science, technology, engineering, and mathematics with real-world applications (Southeast Comprehensive Center, 2012). It is a "meta-discipline – a convergence of science, technology, engineering and math – that offers a student-centered, inquiry-based method of addressing and solving problems" (Southeast Comprehensive Center, 2012, p. 7). The overall purpose is to raise awareness of STEM in society, motivate students, and increase interest in STEM. The anticipation is that an increasing number of students will pursue STEM subjects in college and then careers in STEM fields. Problem-based learning (PBL), Project-based learning, and hands-on discovery/inquiry learning are a few pedagogies advocated for STEM education (Euefueno, 2019). Additionally, the report, *STEM 2026: A Vision for Innovation in STEM Education* (U.S. Department of Education, Office of Innovation and Improvement, 2016) called for "educational experiences that include interdisciplinary approaches to solving grand challenges" (p. ii).

The term STEM representing science, technology, engineering, and mathematics disciplines, arguably, was introduced within the U.S. education system in the early 2000s. The term STEM appears in federal policy within the *America Competes Reauthorization Act of 2010* (2011). Considering the interdisciplinary nature and broader impacts of STEM disciplines, it is possible that if properly designed and implemented, STEM curricula including PBL teaching and learning through hands on activities could be integrated with arts, language arts, reading, and social studies to help students explore the world around them. This interdisciplinary approach to STEM education may lead to higher student motivation in subjects such as science and mathematics and prevent knowledge fragmentation by teaching subjects in isolation (Drake & Burns, 2004). This may be because separation or compartmentalization of subject areas where content is taught discretely during different times of the school day disrupts the learning of many students. Often student comprehension of complex topics increases through an interdisciplinary approach to teaching and learning (Fogarty, 1991). Including interdisciplinary pedagogy is an important aspect of defining STEM education (Drake & Burns, 2004).

STEM Education, Competitive Workforce, and Economic Development

The link between STEM education, a competitive workforce, and economic development is explored, debated and established (Croak, 2018; Hanushek & Kimko, 2000; Hanushek & Wößmann, 2008; Lazio & Ford, Jr., 2019; Oberoi, 2016; World Economic Forum, 2020). There is much discourse in support of K-12 STEM for a competitive workforce by heads of states, legislators, policymakers, industrialists, and business leaders and educators (Kumar, 2019). Croak (2018) in an international analysis, explained a positive link between post-secondary STEM education, human capital and competitiveness, and overall economic development. According to Oberoi (2016), the impact on the economy of introducing STEM at an early age in schools along with academic interventions and support is considerable. K-12 STEM education influences success in post-secondary STEM disciplines, with the subsequent connection to a skilled workforce and economic impact. On the other hand, critics of STEM education would argue, that "focusing on STEM is not enough. Educating young people in these subject areas may ensure they are experts on specific topics, but it does not necessarily create conscientious citizens who can make responsible social and financial decisions" (Billimoria, 2017, n.p.). Though this criticism is aimed at STEM education, it reflects the general state of K-12 education in the U.S.

Challenges to K-12 STEM Education

Though STEM is often portrayed as a priority in K-12 and college settings, the challenge remains on motivating students who repeatedly failed to develop an interest in STEM subjects during their K-12 school years to pursue STEM education in college. If the aptitude and interest for STEM subjects are cultivated in students during the K-12 school years, then the chances of students pursuing STEM degrees in college are remarkably high (Banilower et al., 2018). The National Science Teachers Association (NSTA) has presented position statements for early childhood (NSTA, 2014) and elementary grades (NSTA, 2002) calling for engaging, exciting, and meaningful science learning opportunities for students from age 3 through preschool, and children from elementary levels (K-grade 5) through middle school levels (grades 6-8) respectively. This recommendation is based on research that children "have the capacity for constructing conceptual learning and the ability to use practices for reasoning and inquiry" (NSTA, 2014, p. 1).

However, the results from the National Survey of Science and Mathematics Education (NSSME+) (Banilower et al., 2018) do not paint an encouraging picture of how one of the STEM subjects, science, is taught in U.S. K-12 classrooms. For example, in 37% of the elementary school classes, 30% of the middle school classes, and 31% of the high school classes, students watched while teachers conducted scientific demonstrations. In comparison, 16% of the elementary school classes, 11% of the middle school classes, and 12% of the high school classes, engaged students in hands-on/laboratory activities. Also, only 8% of elementary teachers, 8% of middle school teachers, and 6% of high school teachers involved their students in Project-based learning activities, an essential pedagogy of STEM education aimed at developing critical thinking and problem-solving skills in real-world contexts (Banilower et al., 2018).

The NSSME+ (Banilower et al., 2018) showed that 54% of science teacher professional development offered locally addressed ways of engaging students in hands-on science. However, only 17% of science teacher professional development offered locally addressed ways to integrate student's cultural backgrounds into science teaching. Notably, only 25% of locally provided science teacher workshops addressed building students' confidence in pursuing science/engineering careers. The survey also noted an unfortunate situation, that time spent on science learning in grades K-3 is18 minutes per day, and in grades 4-6 is 27 minutes per day (Banilower et al., 2018). In terms of educational qualifications of science teachers, 3% of elementary teachers hold undergraduate degrees in science/engineering, 1% in science education, and 3% in science/engineering/science education (Banilower et al., 2018). For 65% of elementary science teachers, the route to educator certification is a bachelor's degree, 22% is a master's degree, 11% is post-baccalaureate program (no master's degree) (Banilower et al., 2018).

With regards to equity, there is a disparity in K-12 STEM education. For example, in the U.S., public schools are classified into four quartiles based on the number of students eligible for Free or Reduced-Price Lunch (FRPL) (U.S. Government Accountability Office, 2018). Under the National School Lunch Program, a child whose family's income does not exceed 130% of the federal poverty level is eligible for the free lunch program. Children whose families' incomes are between 130% and 185% of the federal poverty level may receive a reduced-price lunch (U.S. Government Accountability Office, 2018). Also, children in Head Start and Migrant Education programs, children in foster care, and children receiving public service under the Runaway and Homeless Youth Act are eligible for FRPL (U.S. Government Accountability Office, 2018). Based on the U.S. Government Accountability Office (GAO) (2018) report on K-12 education, there is disparity in the offering of high school courses in STEM disciplines between the high (FRPL recipients 75-100%) and low poverty (FRPL recipients 0-24.9%) schools. High-poverty schools are less likely to offer science (i.e., Physics, Chemistry) and mathematics courses that most colleges expect their students to have in high school except for Biology. This is a severe concern when accounting for considerable differences in demographics of high and lowest poverty schools as presented in Table 2 (U.S. Government Accountability Office, 2018).

With respect to students with disabilities, as classified under the U.S. Individuals with Disabilities Act (IDEA) (2004), the situation is not encouraging. The GAO (2018) found a negative association with schools enrolling increasing percent students with disabilities and the likelihood of offering high school courses in Biology, Chemistry, Physics, and AP Science. As Schneiderwind and Johnson (2020) noted "students with disabilities therefore remain underrepresented in STEM fields, and a need exists to help uncover barriers that students with disabilities encounter in STEM laboratories, for example" (n.p.).

Another hindrance to STEM education is the ongoing disparity in per-pupil expenditure, as revealed by the National Survey of Science and Mathematics Education, NSSME+ (Banilower et al., 2018). In the U.S., for schools with students eligible for FRPL, there is inequality in the median amount of dollars spent per pupil between the highest quartile and the lowest quartile in 2018 and 2012. Figure 1 includes the original data by Banilower et al. (2018) and Banilower et al. (2013), adjusted for inflation.

Table 2

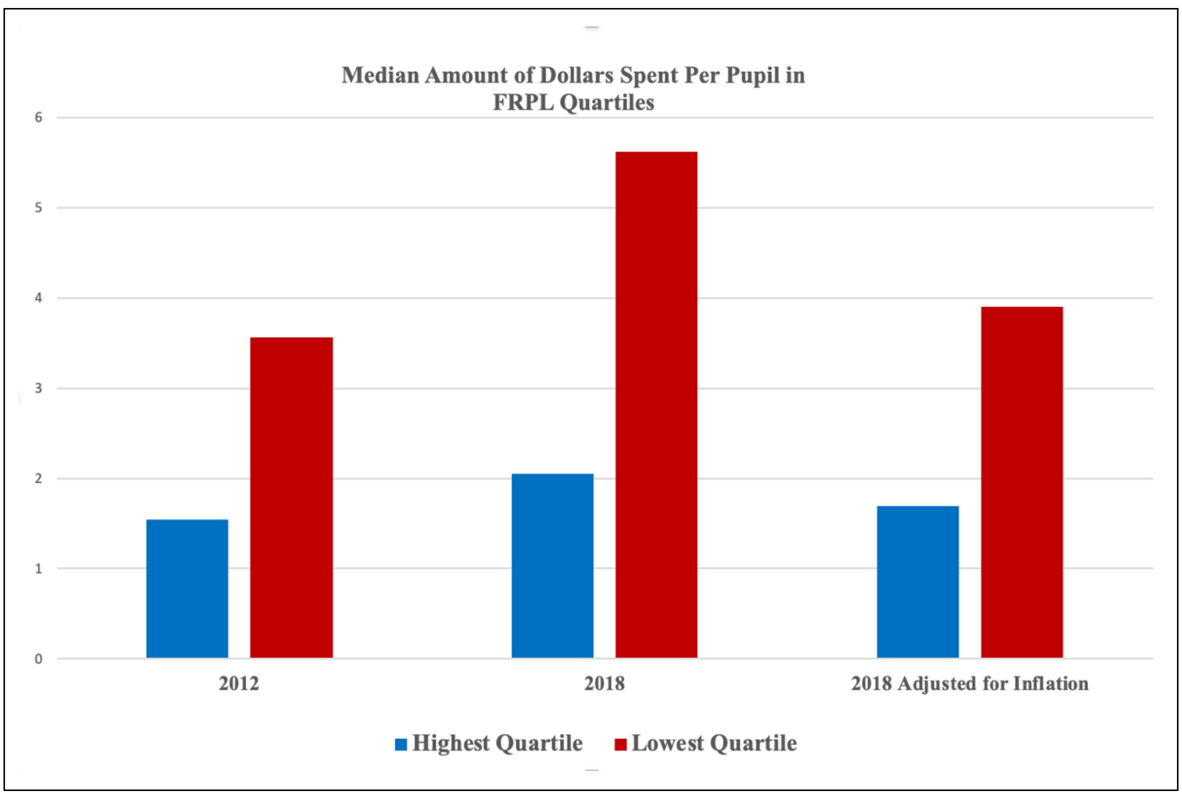
Poverty Level and STEM Course Offerings in School Year 2015-2016

Poverty Level Based on FRPL Eligibility	Demographics	Percent Schools Offering STEM Courses
Highest (FRPL 75-100%)	White 13%, Black 29%, Hispanic 52%, Asian 4%, Other 4%	*Biology 94.35% Chemistry 81.2% Physics 62.5% Advanced Placement Science 69.6% Advanced Placement Math 75.2%
Lowest (FRPL 0-24.9%)	White 71%, Black 6%, Hispanic 11%, Asian 9%, Other 4%	*Biology 97.6% Chemistry 93.5% Physics 89% Advanced Placement Science 88.7% Advanced Placement Math 94.1%

Note: *Trend in Biology seems different compared to other science disciplines.

Figure 1

FRPL Per Pupil Expenditure in Science Adjusted for Inflation



Per pupil expenditure in 2018, in the highest quartile is $2.05 ($1.69) and lowest quartile $5.62 ($3.90) compared to 2012, where it is $1.54 and $3.56, respectively. Percent increase of 33% (10%) in the highest quartile and 58% (10%) in the lowest quartile reveals a large socioeconomic disparity, before adjusting for inflation. To complicate things further, according to NSSME+*,* overall, 52% of classes in the highest quartile with a high proportion of FRPL students are less likely to be taught by teachers with a substantial science background, in terms of having a degree or at least three advanced science courses, compared to classes in the lowest quartile (66%) (Banilower et al., 2018). This is a clear indication of the ongoing socioeconomic divide in science education, a key component discipline of STEM education, needing constructive long-range solutions.

Another challenge to STEM education involves teacher assigned grades, as indicated in a study of a large metropolitan school district in California (Kunnath & Suleiman, 2018). Teacher assigned grades indicated that educators in mid and high poverty schools assigned significantly less "A" grades than low poverty schools (Kunnath & Suleiman, 2018). This study used a survey method to determine the extent to which grading practices and grading influences are used by teachers across subject areas, between poverty levels based on FRPL, when preparing report card grades for students. A possible explanation of why students at mid and high poverty schools are assigned less "A" grades might be that "students of high-poverty schools often come from low-SES households and are less likely to have parents who are actively involved in school, lowering the likelihood of adding pressure on teachers to alter grading practices" (Horvat et al., 2003; Lee & Bowen, 2006 cited in Kunnath & Sulieman, 2018, p. 11). If this is the situation in U.S. K-12 schools, then the goal of education for all, and especially in STEM for All is difficult to reach, perpetuating inequity in the U.S education system. Accordingly, how to address these challenges is a critical question facing U.S. K-12 STEM education. In this context, strategies for addressing these challenges in reforming STEM education as a human resource pipeline for a competitive workforce and economic development are explored.

Strategies for Reforming K-12 STEM Education

It is apparent that in its current state, K-12 STEM education is not a dependable human resource pipeline for any projected competitive STEM workforce and economic development (Banilower et al., 2018). To reform STEM education into a dependable pipeline, the following strategies are essential.

STEM for All

To promote STEM for All, everything possible should be done to promote diversity, and the inclusion of marginalized student groups in STEM education. "Articulate a clear vision for, and long-term commitment to, broadening participation in STEM" of persons with disabilities, women, and under-represented racial/ethnic groups in STEM education (Powell et al., 2018, n.p.; Southeast Comprehensive Center, 2012; Hill & Kumar, 2013; Kumar & Chubin, 2000). Based on poverty levels, race/ethnicity, and disabilities, there is apparent bias in terms of STEM course offerings and the amount of dollars spent per pupil (Banilower et al., 2018). Moreover, most teachers lack training in strategies to integrate students' cultural backgrounds into science pedagogy. Additionally, any bias in teacher assigned grades in STEM classrooms, needs to be addressed without delay. Inequitable, variable, and inconsistent grading practices may negatively affect education for all (Feldman, 2018), especially STEM for All.

These are complex matters that demand committed efforts from the stake holders of K-12 STEM education to find creative solutions towards STEM for All students and promote broad participation. If STEM education is truly a priority in U.S. schools, as touted by U.S. legislatures and leaders in business, then concerted and organized efforts to provide STEM for All should be taken. As the teacher is key to classroom reform, it is critical to emphasize the role of teachers through professional development, retention, and a supportive working environment.

Teacher Training

The most crucial strategy deals with offering appropriate training for teachers in preparation, and teachers currently working in K-12 classrooms, to improve their content and pedagogical knowledge and understanding of STEM education (Kumar & Moffitt, 2022). This will allow teachers to implement meaningful STEM lessons in their K-12 classrooms. The National Commission on Teaching and America's Future (1996) stated that teachers' knowledge and practices are the most significant factors affecting student learning. Therefore, for successful systemic reform in STEM education, it is imperative that classroom teachers are equipped with the knowledge and skills to teach STEM subjects meaningfully to all students. The *National Science Education Standards* (National Research Council, 1996) calls for giving practicing teachers the "same opportunities their students will have to develop understanding" (p. 60) of science, and recommends professional development with more emphasis on "inquiry into teaching and learning; learning science through investigation and inquiry; integration of science and teaching knowledge; etc." (p. 72). Teachers also need assistance in realizing their "blind spots" to create awareness of appropriate teaching and learning strategies for all students. Schools, school districts, and university/college teacher training programs should partner to strengthen the professional development of both teachers in preparation and in service teachers (Schneiderwind & Johnson, 2020).

One consideration is recent deregulation and the subsequent rise in alternate teacher preparation programs, which often replace the requirement of a bachelor's degree in education and the specific discipline (i.e., science and mathematics) and have compromised training in STEM content areas (Perez & Kumar, 2018). Individuals who entered teaching via alternative certification programs often have fewer courses or training hours to complete and are "25% more likely to leave their schools and the profession" (Carver-Thomas & Darling-Hammond, 2017, p. vi). In the name of education reform and politically motivated attempts to sideline university-based teacher training programs, alternative teacher training programs have grown all over the U.S., offering less comprehensive, and inadequately regulated teacher training where STEM subjects, especially science, are often not a priority (Ildiko & Berliner, 2002).

Teacher Retention

Increasing teacher retention, particularly in STEM subjects is critical to the successful implementation of STEM education. Teacher attrition is not only an educational crisis, but it also has severe economic and human resource implications. For example, from an economic standpoint, it costs approximately $21,000 USD to replace each teacher in an urban school, therefore reducing attrition in half would save $10,500 USD per urban school teacher (Carver-Thomas & Darling-Hammond, 2017). During the 2015-2016 school year, 40 U.S. states and the District of Columbia (D.C.) reported a teacher shortage in science, and forty-two states and D.C. reported a teacher shortage in mathematics (Sutcher et al., 2016). Disparities in the teacher labor market change from U.S. school district to district in critical shortage subject areas (Sutcher et al., 2016). High teacher attrition also contributes to low student achievement. It is disheartening that an economically prosperous nation such as the U.S. does not want to pay its teachers a competitive market salary. However, teachers are expected to train students into a STEM competent workforce. Senior teachers' turnover rate at the top of their district salary (average $78,000 USD) schedules is 31% lower than teachers with top district salaries below $60,000 USD (Carver-Thomas & Darling-Hammond, 2017), indicating that higher salaries may support a lower attrition rate. Overall, the predicted turnover rate of mathematics and science teachers is 37% higher than elementary teachers (Carver-Thomas & Darling-Hammond, 2017). So, if the U.S. wants to support K-12 STEM education, teachers need to be compensated adequately.

Supportive Environment for Teachers

An analysis of a teacher follow-up survey by the Learning Policy Institute (2017) showed that among several other reasons, 21% of teachers leave the field because of dissatisfaction with the administration, and 25% leave teaching positions due to dissatisfaction with school assessment policies. Schools where "principals generally describe their leadership responsibilities as facilitators, collaborators, team leaders, or leaders of leaders" have low teacher attrition rates (Learning Policy Institute, 2017, p. 2). To reform STEM education, it is imperative that the school administration provides a supportive environment for teachers to utilize their abilities to lead and inspire students to learn (Kumar & Chubin, 2000). School districts throughout the U.S. should not only aspire towards improving the work environment for teachers, but also for school administrators (i.e., principals) to enhance teacher retention in K-12 education (Carver-Thomas & Darling-Hammond, 2017). In a study limited to a large school district in Arizona, Sulit (2020) found that the distributive leadership framework significantly increased teacher retention in elementary and middle grades. This indicates that there is still hope for improving teacher retention. Suitable policies are needed in this area in K-12 education, not only for the sake of STEM disciplines, but also for all disciplines since K-12 education is an extremely critical human development process in a child's life.

Concluding Thoughts

STEM education is a human enterprise needing human, material, and fiscal resources to be successful. In this context, the classroom teacher is a significant catalyst for transforming K-12 STEM education as a dependable pipeline for a competitive workforce and economic development. The strategies discussed will work, provided socioeconomic disparities that impact public education in the U.S. are adequately addressed and resolved. It is extremely important that leaders of industries and businesses collaborate with legislators, policymakers, school administrators, classroom teachers, and parents to transform K-12 STEM education as a dependable human resource pipeline for a competitive workforce and economic development.

The authors received no financial support for the research, authorship, and/or publication of this manuscript.

David Devraj Kumar (david@fau.edu) is a Professor of Science Education and Founding Director of the STEM Education Laboratory at Florida Atlantic University, and a former Visiting Fellow in Governance Studies at the Brookings Institution. His research interests include STEM teaching and learning, computer supported problem-based learning, evaluation, and policy. He has served as the Principal/Co-Principal Investigator of projects funded by the National Science Foundation, U.S. Department of Education, and the Ewing Marion Kauffman Foundation. In recognition of his research and scholarly contributions, he was awarded the Sir Ron Nyholm Education Prize by the Royal Society of Chemistry and elected as a Fellow by the International Academy of Education, and the American Association for the Advancement of Science.

Susannah L. Brown (sbrow118@fau.edu) Professor, Art Education, College of Education, Department of Curriculum and Instruction, works with all ages, PreK-adults, to inspire academic excellence and a commitment to the community through arts engagement for over thirty years. Cross disciplinary expertise in the arts and education along with teaching experience with diverse learners, propels Dr. Brown forward as an engaged faculty member and mentor for her students.

References

America Competes Reauthorization Act of 2010, Pub. L. No. 111-358, 124 Stat. 3982 (2011). https://www.govinfo.gov/content/pkg/PLAW-111publ358/html/PLAW-111publ358.htm

Banilower, E. R., Smith, P. S., Malzahn, K. A., Plumley, C. L., Gordon, E. M., & Hayes, M. L. (2018). *Report of the 2018 NSSME+, December 2018*. Horizon Research, Inc.

Banilower, E. R., Smith, P. S., Weiss, I. R., Malzahn, K. A., Campbell, K. M., & Weis, A. M. (2013). *Report of the 2012 national survey of science and mathematics education*. Horizon Research, Inc.

Billimoria, J. (2017, May 4). *Is STEM education all it's cracked up to be?* World Economic Forum. https://www.weforum.org/agenda/2017/05/is-stem-education-all-it-s-cracked-up-to-be/

Carver-Thomas, D. & Darling-Hammond, L. (2017). *Teacher turnover: Why it matters and what we can do about it*. Learning Policy Institute.

Committee on STEM Education of the National Science and Technology Council. (2018). *Charting a course for success: America's strategy for STEM education.* U.S. Government.

Croak, M., (2018). *The effects of STEM education on economic growth.* (Publication number: 1705). [Honors Theses, Union College]. https://digitalworks.union.edu/theses/1705

Drake, S. & Burns, R. (2004). *Meeting standards through integrated curriculum.* Association for Supervision and Curriculum Development.

Euefueno, W. D. (2019). Project-/problem-based learning in STEM: Impacts on student learning. *Technology & Engineering Teacher*, *78*(8), 8-12.

Feldman, J. (2018, December). Correcting variable grading policies can transform an entire school. *Principal Leadership, 19*. https://www.nassp.org/publication/principal-leadership/volume-19-2018-2019/principal-leadership-december-2018/inequity-in-grading/

Fogarty, R. (1991). Ten ways to integrate curriculum. *Educational leadership*, *49*(2), 61-65.

Hanushek, E. A. & Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *American Economic Review, 90*(5), 1184-1208. https://doi.org/10.1257/aer.90.5.1184

Hanushek, E. A. & W¨oßmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature, 46*(3), 607-668.

Hill, J. & Kumar, D. D. (2013). Challenges for chemical education: Implementing the 'chemistry for all' vision. *The Chemist, 86*(2), 27-32.

Horvat, E., Weininger, E., & Lareau, A. (2003). From social ties to social capital: Class differences in the relations between schools and parent networks. *American Educational Research Journal, 40*(2), 319-351. https://doi.org/10.3102/00028312040002319

Ildiko, L. & Berliner, D. (2002). The Effectiveness of "Teach for America" and Other Under-certified Teachers. *Education Policy Analysis Archives.* https://doi.org/10. 10.14507/epaa.v10n37.2002.

Individuals with Disabilities Education Act, 20 U.S.C. § 1400 (2004). https://sites.ed.gov/idea/statute-chapter-33/subchapter-i/1400

Kumar, D. D. & Moffitt, S. (2022). Understanding the STEM teacher pipeline: Early findings from a review of selected NSF Noyce program evaluations. *Voices for Educational Equity, 18*(2), 50-57.

Kumar, D. D. (2019, November). *Missing factors in economics based arguments leveraging STEM education for a competitive workforce*. A paper presented at the Critical Questions in Education Symposium, The Academy for Educational Studies, Chicago, IL.

Kumar, D. D. & Chubin, D. E. (Eds.) (2000). *Science, technology, & society: A sourcebook on research and practice*. Kluwer Academic/Plenum Publishers.

Kunnath, J. P. & Sulieman, M. (2018). Examining the role of poverty in teacher grading decisions. *Journal of Leadership, Equity, and Research, 4*(2), 1-22. https://journals.sfu.ca/cvj/index.php/cvj/article/view/43

Lazio, R. & Ford, Jr., H. (2019, October). *The economic imperative of reforming STEM education*. Scientific American. https://www.scientificamerican.com/article/the-economic-imperative-of-reforming-stem-education/

Learning Policy Institute. (2017, February). The role of principals in addressing teacher shortages. *Research Brief*. Author. https://learningpolicyinstitute.org/product/role-principals-addressing-teacher-shortages-brief

Lee, J. & Bowen, N. K. (2006). Parent involvement, cultural capital, and the achievement gap among elementary school children. *American Educational Research Journal, 43*(2), 193-218. https://www.jstor.org/stable/3699418

National Commission on Teaching and America's Future. (1996). *What matters most: Teaching for America's future*. Author. https://www.teachingquality.org/wpcontent/uploads/2018/04/What_Matters_Most.pdf

National Research Council. (1996). *National Science Education Standards*. National Academy Press. https://nap.nationalacademies.org/catalog/4962/national-science-education-standards

National Science Teachers Association. (2014). *NSTA position statement: Early childhood science education*. Author. https://www.nsta.org/nstas-official-positions/early-childhood-science-education

National Science Teachers Association. (2002). *NSTA position statement: Elementary science*. Author. https://www.nsta.org/nstas-official-positions/elementary-school-science

Noonan, R. (2017). *STEM jobs: 2017 update (ESA issue brief #02-17)*. U.S. Department of Commerce, and Statistics Administration, Office of the Chief Economist. https://www.commerce.gov/sites/default/files/migrated/reports/stem-jobs-2017-update.pdf

Oberoi, S. (2016). *The economic impact of early exposure to STEM education*. The Committee for Economic Development of The Conference Board. https://www.ced.org/blog/entry/the-economic-impact-of-early-exposure-to-stem-education

Perez, K. & Kumar, D. D. (2018). STEM professional development policies in the United States: Trends & Issues. In M. Shelley & A. Kiray (Eds.), *Research highlights in STEM education*. ISRES Publishing.

Powell, A., Nielsen, N., Butler, M., Buxton, C., Johnson, O., Ketterlin-Geller, L., & McCulloch, C. (2018). *Creating inclusive PreK–12 STEM learning environments*. Education Development Center. https://files.eric.ed.gov/fulltext/ED590489.pdf

Schneiderwind, J. & Johnson, J. (2020, July 27). Why are students with disabilities so invisible in STEM education? *Education Week*. https://www.edweek.org/education/opinion-why-are-students-with-disabilities-so-invisible-in-stem-education/2020/07

Southeast Comprehensive Center. (2012). *Engaging diverse learners through the provision of STEM education opportunities. Briefing Paper*. Author. https://sedl.org/secc/resources/briefs/diverse_learners_STEM/Diverse_Learners_through STEM.pdf

Sulit, A. (2020). Leadership can't stand alone: Why school districts need policy to increase teacher retention. *eJournal of Education Policy*, *21*(2). https://doi.org/10.37803/ejepF2008

Sutcher, L., Darling-Hammond, L., & Carver-Thomas, D. (2016). *A coming crisis in teaching? Teacher supply, demand, and shortages in the U.S*. Learning Policy Institute. https://doi.org/10.54300/247.242

U.S. Department of Education, Office of Innovation and Improvement. (2016). *STEM 2026: A vision for innovation in STEM education*. Author. https://oese.ed.gov/files/2016/09/AIR-STEM2026_Report_2016.pdf

U.S. Department of Labor, Bureau of Labor Statistics. (2017). *STEM occupations: Past, present, and future*. Author. https://www.bls.gov/spotlight/2017/science-technology-engineering-and-mathematics-stem-occupations-past-present-and-future/pdf/science-technology-engineering-and-mathematics-stem-occupations-past-present-and-future.pdf

U.S. Government Accountability Office. (2018). *K-12 education. Public high schools with more students in poverty and smaller schools provide fewer academic offerings to prepare for college*. Author. https://www.gao.gov/products/gao-19-8

World Economic Forum. (2020). *Schools of the future defining new models of education for the fourth industrial revolution*. Author. https://www.weforum.org/reports/schools-of-the-future-defining-new-models-of-education-for-the-fourth-industrial-revolution/

ELECTRONIC JOURNAL FOR RESEARCH
IN SCIENCE & MATHEMATICS EDUCATION
VOL. 27, NO. 3, 75-95

Stakeholders' Conceptions of STEM and Elementary STEM Clubs Within a Community-University Partnership

Richard Carlos L. Velasco
University of Oklahoma

Rebecca Hite
Texas Tech University

ABSTRACT

Informal community-based science, technology, engineering, and mathematics (STEM) clubs provide rich informal learning environments that help elementary-aged students develop STEM knowledge and skills while fostering their initial and continued interests in STEM. This phenomenographical case study sought to interpret stakeholders' (five university personnel, two club facilitators, one teacher, one parent, and three elementary students) conceptions of STEM and STEM clubs involved in a community-university partnership in an afterschool elementary STEM club at a community center. Phenomenographic analyses produced three hierarchical categories in stakeholders' conceptions of STEM: an indifference towards STEM, viewing STEM as a holistic discipline, and STEM as applicable and useful in life. Among stakeholders' conceptions of STEM clubs, four hierarchical categories described clubs as a non-STEM related space, a means to promote STEM, provide STEM learning, and an additional site (apart from school) to produce STEM knowledge, skills, and enjoyment. Findings suggest that community and university stakeholders held varying conceptions of the purpose of STEM, with the strongest disagreement in how informal STEM clubs should be structured. Stakeholders nonetheless agreed that STEM clubs were vital resources to promote STEM and enhance STEM-related life and soft skills.

Keywords: community-university partnerships, elementary education, informal education, phenomenography, STEM clubs

Introduction

The National Science Foundation, among other national level educational organizations, has lauded that club-based science, technology, engineering, and mathematics (STEM) *informal* (per Eshach, 2007) experiences enhance STEM learning and affect for K-12 students. The National Research Council's (NRC, 2015) report on *Identifying and Supporting Productive STEM Programs in Out-of-School Settings* found that K-12 students spend only 20% of their waking hours in school, where the remaining 80% spent outside of school could include participating in cooperative STEM learning activities (p. 8). Further, the NRC stated that STEM clubs can contribute to students' success in STEM by providing hands-on learning activities with peers to develop both cognitive (content knowledge, academic) and non-cognitive (soft skills, affective) skills, supplementing STEM learning and enjoyment unmet due to the constraints (e.g., time, curriculum, space) of the formal K-12 STEM classroom. Related research affirms the NRC evidence that students who participate in STEM clubs have improved achievement in cognitive and non-cognitive domains (Blanchard et al., 2017; Hite et al., 2018;Hite & White, 2019, 2021; Sahin, 2014) and persistence in STEM (Gottfried & Williams, 2013). Given that STEM interests can be cultivated as young as the elementary level (Bybee & Fuchs, 2006), access to STEM clubs among primary students is desired among many communities looking to provide younger students access to STEM outside of the classroom (DeJarnette, 2012), where science and math subjects receive the least amount of instructional time (Blank 2012; Lavy, 2010). Since these activities occur outside of school, community partnerships (with local universities) can help design and implement STEM clubs to improve outcomes for K-12 students, especially those that are under resourced and may not have access to rich, out-of-school science experiences (Duodu et al., 2017).

This paper describes a community-university partnership (CUP) formed between local K-12 educators and a research-oriented university with a mutual aim to improve elementary STEM education through afterschool STEM clubs. We define this CUP as "collaborations between *community* organizations and institutions of higher learning for the purpose of achieving an identified social change goal through *community*-engaged scholarship that ensures mutual benefit for the community organization and participating students" (Curwood et al., 2011, p. 16). The STEM club in this case study was developed internally (by university faculty and staff), in consultation with technical assistance from the research literature, and with input from the community partners (i.e., day of week, age group of interest). This relationship is a step in the right direction, as per the NRC report, "research is needed to better specify and understand the ways in which learning develops across formal and informal settings, [especially in] leveraging community resources and partnerships" (NRC, 2015, p. 29).

In order for the CUP to reach its mutual goals, it is vital that the community stakeholders understand the premise of the CUP and its intended utility (Curwood et al., 2011). Doing so not only provides greater input to the design and implementation process (of STEM Clubs), but also to amplify the importance of STEM for elementary aged learners in the community. As such, a community's understanding in or value of STEM clubs or the premise of STEM itself is necessary. Therefore, the purpose of this study was to explore community stakeholders' conceptions of STEM and STEM clubs. By exploring the understandings and expectations of the CUP STEM community, we may foster more productive relationships between stakeholders and these informal STEM programs. The outcomes of this study will be used to inform current and future STEM club programming so we may better leverage CUP stakeholder input and resources to enhance our local pipeline of STEM savvy elementary students.

Literature Review

Definitions and conceptions of STEM are diverse and vague (Bybee, 2010), influencing researchers to explore a more definitive answer to *What is STEM?* There is a growing body of literature that showcases conceptualizations of STEM from specific groups of the STEM ecosystem. For instance, Breiner et al. (2012) found that university faculty and staff have differing conceptions on STEM, depending on the relevance and impact that STEM has on their personal and professional lives. Further, they described how university personnel, especially non-STEM faculty, were more indifferent towards STEM. Most faculty—STEM faculty included—viewed STEM in its individualized disciplines rather than as a holistic whole. Classroom teachers also hold similar conceptions of STEM but have shown to lean more towards integration when immersed in effective STEM education professional development (Ring et al., 2017), or compelled by integrated standards (NGSS Lead States, 2013). Research has shown that afterschool STEM club facilitators—who may not necessarily be certified classroom teachers—possess strong STEM identities due to the authentic, real-world STEM activities they plan and implement in STEM clubs (Aslam et al., 2018). Much of their conceptions on STEM clubs involve the development of students' technical STEM skills and helping students to ground theoretical knowledge in real life applications. Thus facilitators with passion for and commitment to the STEM club is fundamental to its long term success (Blanchard et al., 2017). Regarding facilitation, Davis et al. (2023) concluded from a systematic literature review of studies on STEM clubs that

> the literature highlights that STEM clubs should be facilitated in a way that is driven by student interest, moves outside of the traditional teacher role, and nurtures in participants the ability to enact peer teaching roles or consider being a possible future facilitator. STEM clubs offer facilitators more flexibility, creativity, and innovation in their teaching than is possible in a more traditional classroom context. (p. 11)

Parents are generally unsure about what STEM means or entails, but nonetheless see the importance and value in learning STEM for their children's future careers (Hernandez et al., 2016). Students' conceptions of STEM, on the other hand, are dependent upon the exposure they receive from parents (Plasman et al., 2021; Tay et al., 2018), formal learning experiences in the K-12 classroom (Mullet et al., 2018), and in the community from informal learning opportunities (Afterschool Alliance, 2015). While a concrete definition has yet to be agreed upon among both researcher and practitioner groups (Radloff & Guzey, 2016), considering how all those involved in the STEM ecosystem (inclusive of parents and students) conceptualize STEM would paint a broader picture of what informal STEM learning is or should be, fostering greater understandings of STEM.

One way to capture community stakeholders' conceptions of STEM are through non-compulsory (compared to formal or school) involvement in STEM, like informal, afterschool community-based STEM clubs. Afterschool STEM clubs have proven to be effective spaces for K-12 students to learn and engage in STEM skills and knowledge not typically learned in the formal classroom setting (Afterschool Alliance, 2015). STEM clubs in the afterschool setting provide access and exposure to STEM students that build critical thinking and problem-solving skills, as well as enhance interest and enjoyment in STEM especially at the elementary level (Ching et al., 2019; Sahin et al., 2014).

Furthermore, research has shown that engaging in STEM clubs involved in a CUP, in particular, have a myriad of benefits for all those involved in the partnership (Hite et al., 2020, 2023; Foster et al., 2010). Multiple studies have showcased community stakeholders—teachers, parents, and students—as well as university personnel having positive outcomes in bilateral learning and understanding of STEM (see Allen et al., 2019; Hite & White, 2022; Playton et al., 2021, 2023; NRC, 2015; Tay et al., 2018; Toma & Greca, 2018). Thus, we find it imperative to study and understand community stakeholders' conceptions of STEM and STEM clubs, doing so through a theoretical framing that permits interpretation on how individuals develop meanings and variation from their understanding (of STEM) and experiences (of/in STEM clubs).

Theoretical Framework

To explore varying understandings and experiences among a group, this study employed the theoretical aspects of phenomenography. Phenomenography initially emerged as an empirical rather than a theoretical or philosophical tradition (Marton, 1981), and was initially viewed solely as a methodological practice (Åkerlind, 2012). However, since phenomenology was first established, Marton (1986) has clarified that phenomenography also undertakes theoretical and ontological perspectives as it provides a model to answer questions about thinking and learning. Since then, phenomenography has been used in various other studies as theoretical or conceptual lenses in addition to a qualitative research design (see Andretta, 2007; Cope, 2004; Ornek, 2008). In phenomenography, learning can be viewed in two different lenses, as first-order and second-order perspectives (Marton, 1986). In a first-order perspective, learning is viewed from the researcher's perspective, specifically how the phenomenon of study is related to their worldview and their understanding of reality. Whereas, learning from a second-order perspective is centrally focused on the ways the participants' experiences (of a phenomenon) mediate their understanding and conceptions (of said phenomenon). For instance, research in best practices on learning STEM is largely viewed from a first-order perspective, whereas studies that examine the ways in which participants experience STEM learning is indicative of the second-order perspective (e.g., Gandhi-Lee et al., 2017; Mullet et al., 2018). This second-order perspective is useful to CUP research to ensure the voices of non-research stakeholders are duly represented.

Perhaps the most significant tenet of phenomenography is its non-dualistic ontology, meaning that a participant's experienced world is neither constructed nor imposed on by the participant, instead it theoretically exists as an internal relationship between the participant's understanding of the phenomena and their experiences with the phenomena (Marton & Booth, 1997). Given the different ways of experiencing a phenomenon, this theory permits modeling of experiences, among very different people, within the same phenomenon. In using phenomenology as a theoretical lens to undergird a study, "the researcher aims to constitute not just a set of different meanings, but a logically inclusive structure relating the different meanings" (Åkerlind, 2012, p. 323). Using phenomenology in this study allowed us to capture the varying definitions and divergent conceptualizations of STEM among different stakeholder groups involved in STEM education. This theoretical perspective compensates for assumptions made in the aforementioned research that all stakeholders involved in STEM clubs share a common understanding of STEM and expectations for out-of-school STEM learning. Thus, a dearth of research remains on how various CUP (community and university) stakeholders conceive of the purpose of STEM and envision students' participation in such clubs.

Without knowledge of how all stakeholders conceptualize STEM and STEM clubs (phenomena of interest), CUP STEM clubs will be unable to reach their full potential in meeting the mutual aims of the university and community in bolstering K-12 STEM learning. In that regard, understanding how CUP stakeholders conceptualize STEM learning via experiences in and conceptions of CUP-based STEM clubs could help inform best practices and improve informal STEM learning spaces for students in the community. Guided by this research approach of phenomenography in the context of this study, this study addresses the following research question: *How do stakeholders involved in a CUP-based elementary STEM club conceptualize STEM and afterschool community STEM clubs?*

Method

Given the duality of phenomenography as a theoretical framework and methodological approach, this study utilized phenomenography as method to "produce an objective, qualitative description to represent the way that individuals perceive reality" (Alsop & Tompsett, 2006, p. 245). Qualitative accounts of stakeholders' conceptions of the phenomenon of STEM and STEM clubs were examined collectively from the sets of participants (stakeholder groups), as opposed to analyzing data from individuals. These accounts, taken from the set of participants, are then organized into what are known as *categories of description*, the primary outcomes of phenomenographic research. While variations exist in the extent to which categories of description are organized (Åkerlind, 2012), the process is both iterative and comparative. Multiple rounds of sorting and grouping are necessary, in addition to comparisons between various participant accounts, as well as between distinct categories of description themselves. Furthermore, a significant premise of these categories is that they are structured in a logical manner, typically hierarchically. This structured and logical set of organized categories form a field that is known as an *outcome space*, wherein "the outcomes represent the full range of possible ways of experiencing the phenomenon in question, at this particular point in time, for the population represented by the sample group collectively" (Åkerlind, 2012, p. 323).

Elementary STEM Club Framework and Context

The conceptual framework for the elementary STEM club in this study was guided by the NRC report's (2015, p. 15) three factors that foster productive STEM club programs: 1) Productive programs engage young people intellectually, socially, and emotionally (e.g. first-hand experiences with phenomena and materials, engaging students STEM practices, and establishing a supportive learning community); 2) Productive programs respond to young people's interests, experiences, and cultural practices (e.g., position STEM as socially meaningful and culturally relevant, support collaboration, leadership, and ownership of STEM learning where staff are co-investigators and learners alongside young people); and 3) Productive programs connect STEM learning in out-of-school, school, home, and other settings (e.g., connect learning experiences across settings, leverage community resources and partnerships, and actively broker additional STEM learning opportunities).

The STEM Club for this present study consisted of university (i.e., faculty, staff) and community (i.e., parents, teachers, students) stakeholders of a local STEM club established through a CUP at a large southwestern research university. The CUP-based STEM club in this research is one among 10-15 active STEM clubs established and led by the university. This STEM club takes place at a local community center, which is unique because most STEM clubs occur at the school location and are not at the elementary level in this community. Participating students mirror the demographics of the community center, as predominantly Hispanic and classified as low socioeconomic status. Approximately twelve elementary students (grades K to 5, four males and eight females) participated in the STEM club, which met once a week for the duration of one typical calendar school year (approximately thirty-six weeks). Activities and content of the STEM club included weather, probability, and algebraic logic. Mobile tablets (iPads) were also used and incorporated into these activities at least once a month.

Participants

Selection of participants in phenomenography research was purposive in that the approach sought to glean participants' conceptions of a phenomenon; in this particular context, defining STEM and conceptualizing STEM clubs. Some phenomenographic researchers suggest a sample size between ten and thirty participants (Mullet et al., 2018; Ornek, 2008). However, other studies have indicated a variation of small and large sample sizes (Gandhi-Lee et al., 2017; Limburg, 2008; Velasco & Hite, 2022).

In total, twelve participants were recruited for this study. As this study sought to analyze the collective conceptions of STEM learning via a community STEM club, it was necessary to sample various participants who encompassed the CUP of this STEM club. The twelve participants in this study consisted of five stakeholders from the community—three students (S1, S2, and S3), one parent (P), and one elementary classroom STEM teacher (T)—and seven stakeholders from the university—staff and faculty personnel (UP1, UP2, UP3, UP4, and UP5) that included the two STEM club facilitators (CF1 and CF2) who were not current classroom teachers. Notably, CF1 had been a mathematics teacher, nationally board certified in early adolescent mathematics and received national recognition for excellence in K-6 mathematics teaching. CF2 held no prior or current teaching credentials. Aside from two university faculty and the elementary STEM teacher, all participants in this study were directly affiliated with the CUP STEM club of study, meaning sampled students were participants in the STEM club, the parent participant was a parent of a STEM club student, community center facilitators assisted with the STEM club, and the three university faculty coordinated and established this specific STEM club as examples. Table 1 describes demographics of participants in this study.

Table 1

Community Stakeholder Participant (n=12) Demographics

Stakeholder	Sex	Ethnicity	Notes
University Personnel (UP)			
UP1*	M	White	Associated with CUP STEM club in this study
UP2**	F	White	Not associated with CUP STEM club in this study
UP3**	F	White	Associated with CUP STEM club in this study
UP4*	F	White	Not associated with CUP STEM club in this study
UP5*	F	White	Associated with CUP STEM club in this study
Club Facilitator (CF)			
CF1	M	Pacific Islander	Facilitator of STEM club and university researcher
CF2	F	Hispanic	Community center overseer of CUP STEM Club
Teacher (T)			
T1	F	Hispanic	Elementary STEM teacher with a focus on science
Parent (P)			
P1	F	Hispanic	Parent of S1, child participant in CUP STEM club
Student (S)			
S1	M	Hispanic	3rd grade student participant in CUP STEM club
S2	F	Hispanic	3rd grade student participant in CUP STEM club
S3	F	Hispanic	4th grade student participant in CUP STEM club

Note. M = male. F = female.
*University faculty
**University staff

Data Sources

The primary source of data for this study were one-time, in-depth, semi-structured interviews with each of the participants. All researchers (i.e., authors of this article) participated in the creation of the interview items, as they related to the tenets of phenomenography in terms of seeking participants' conceptions of STEM learning via informal STEM clubs (see Appendix A for protocols). Literature from the *Informal Learning Report* (National Academy of Engineering & National Research Council, 2014) and Funds of Knowledge framework (Moll et al., 1992) were consulted in protocol development to inform question design that related to STEM understandings and the relevancy of STEM, respectively. Interview items queried experiences along the same themes of a community STEM club and STEM learning in general across all interviews, although interviews were slightly modified based upon the interviewee. Interview questions were simplified to ensure that items were comprehendible for the students, but nonetheless followed the same line of inquiry regarding their conceptions of the STEM club they were participating in and their experience in STEM learning in general. The student participants were interviewed for about fifteen minutes, and the adult participants were interviewed for about thirty minutes. All participants were interviewed about one month after the start of the STEM club, and all interviews were audio-recorded. The online interview transcription application, Otter (2020), was used to transcribe all interviews. Transcriptions were then audio reviewed thoroughly to verify interview segments that were erroneously transcribed from the software.

Analysis

There is not one prescribed technique in analyzing phenomenographic research, as different phenomenographers employ a variation of frameworks from one study to the next (Åkerlind, 2012). Retrospectively, Marton (1986) argued that there are no specific algorithms to discover conceptions of a phenomenon, rather just a proposed set of guidelines to employ when evaluating participants' understandings of a phenomenon. His recommended guidelines are conducted in the following phases: (1) selection of utterances based on criteria relevance—the group of utterances formed from this selection is referred to as the *pool of meanings*; (2) interpretation of the pool of meanings; (3) sorting and arranging utterances into categories of description; (4) differentiating between and refining categories; and finally, (5) defining categories with supporting quotes. Marton's suggested guidelines for phenomenographic analysis was adopted and followed stepwise as the analytical framework for this study.

The first phase of analysis involved the selection of utterances from each interview transcript regarding participants' conceptions of STEM and STEM clubs. Utterances were typically one to two sentences in length and included segments or partial phrases stemming from sentences. This was to ensure that sentences with multiple meanings could be analyzed and represented in the data set as such. The extracted utterances from the transcripts were grouped together, without any stakeholder designations, in a separate file forming a pool of meanings (of 240 utterances) for preliminary coding.

The second phase involved interpretation of the pool of meanings, which also consisted of writing memos to find similarities and differences among utterances, both as a whole (dataset) and as they related to the transcripts. Analysis proceeded to the third phase to sort and rearrange utterances into categories of description. Utterances were first grouped into preliminary categories across the pool of meanings based on similarities. Moreover, analysis of collective meanings across the data was a focal point when grouping the utterances into the preliminary categories. For example, the utterances 'I think that imparting knowledge about how STEM is applied in the world today is integral in a STEM club' and 'I feel like a stem club should teach them things that they're not necessarily going to experience, whether its whole to kind of expand their learning and their knowledge' were sorted into a preliminary category called 'STEM exposure.'

If utterances diverted or differed from those within existing preliminary categories, a new category was created. In a few instances, utterances were dropped completely due to the irrelevance to the data as a whole. Twenty-eight utterances were uncategorized and therefore dropped, resulting in the final remaining 212 utterances sorted into seventeen preliminary categories (see Appendix B). The fourth phase of the analysis involved differentiation and refinement and involved a more focused view on relationships among and between the preliminary categories. As a result, some preliminary categories were combined, consolidated, or collapsed depending on the collective meanings across the utterances, resulting in seven hierarchical categories. Finally, the fifth and final phase of the analysis involved assigning definitions and supporting quotes to each of the core categories. Sub-category coding was performed to provide greater visualization of the utterances associated with the larger categories. Stakeholder designations were added back to the coded utterances for data visualization purposes. The tiered categories and their explanations, with supporting quotes, are provided in the results section.

To ensure trustworthiness of the analyses in qualitative phenomenographic research, communicative checks (Kvale, 1996) were carried out to verify the research methods and interpretations of the data with other members of the research community. As such, the first author of this study employed bracketing (as explained previously) by examining and evaluating presuppositions of the phenomenon of study (i.e., what is STEM and STEM clubs), documented these processes through audits (see Appendix A and Table 1), and verified categories of description with the co-author, a member of this research community. Pragmatic validity checks were also employed to ensure the usefulness of the outcomes of the study (Kvale, 1996). As such, this study also serves as a response to NRC's (2014) call to improve STEM learning in informal learning spaces by involving all stakeholders' conceptions (rather than single groups) in CUP-based STEM clubs. Regarding trustworthiness in interview analysis, authors combined data, stripping off the stakeholder designations in creating the pool of meanings and developing categories. This process was repeated until final categories were established. Researchers (i.e., authors of this article) also met monthly to discuss data and analytical methods, as well as shared data over a secure online database. Last, the second author double coded the data set analyzed by the first author for full agreement.

Results

Stakeholders' Conceptions of STEM

Completion of analysis resulted in an outcome space of three hierarchical core categories of description (fifty-one utterances), which were labeled as follows: Category (1) *Indifference Towards STEM* with fourteen utterances; Category (2) *STEM as a Holistic Discipline* with eleven utterances; and Category (3) *Applicability and Usefulness of STEM* with twenty-six utterances. Table 2 defines these core categories of description for stakeholders' conceptions of STEM, supported by examples of utterances extracted from the data across this outcome space.

Table 2

Core Categories of Description for Stakeholders' Conceptions of STEM

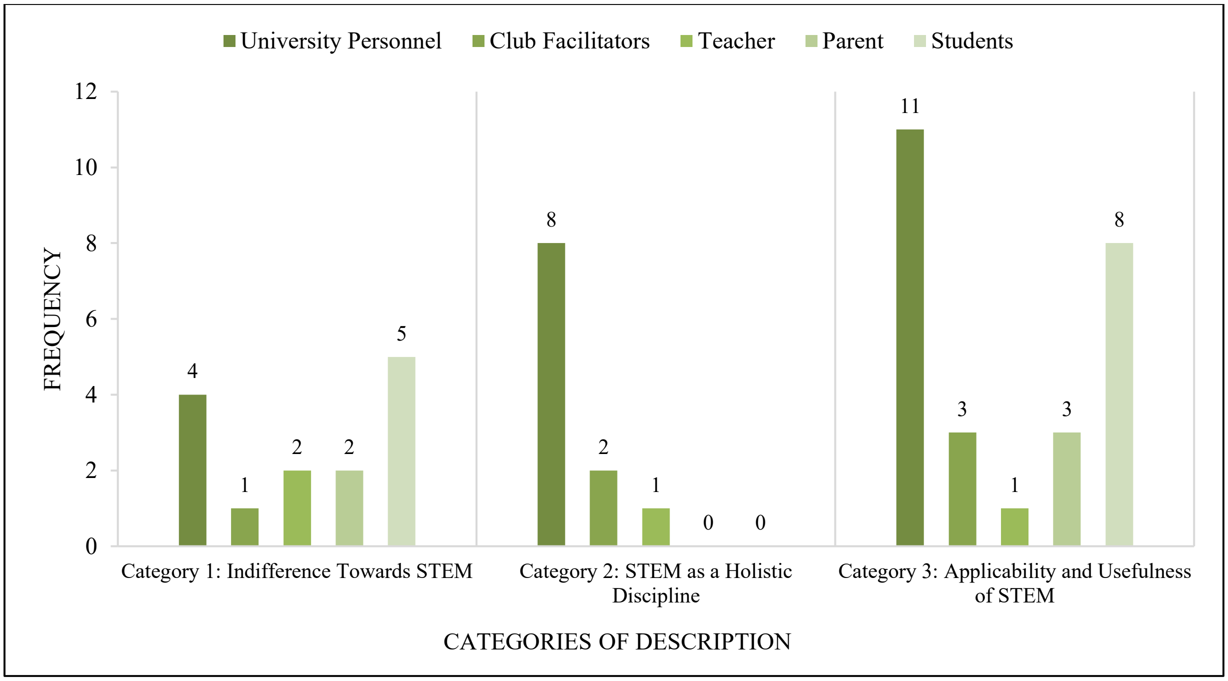
Core Category	Definition	Examples of supporting quotes (utterances)
Category 1 (*n* = 14) Indifference Towards STEM	Stakeholders had no knowledge of understanding of STEM or considered STEM in its individual disciplines or merely an acronym.	UP1: "…in STEM, I feel you need to be in one of the four disciplines of science, technology, engineering and mathematics." T: "I know what it stands for: science, technology, engineering and math and just having kids to be able to apply those four things in."
Category 2 (*n* = 11) STEM as a Holistic Discipline	Stakeholders' conceptions of STEM were that the disciplines were integrated or interrelated in some aspect.	UP2: "I think it's more integrated. I think that they all overlap. I think it's hard to do any of them without the other." CF1: "…STEM implementation is addressing a problem through the use of interdisciplinary skills in STEM."
Category 3 (*n* = 26) Applicability and Usefulness of STEM	Stakeholders conceptualized STEM as applicable and useful in real-life situations for home, school, and for the future.	UP4: "I would stress STEM is everywhere – and everyone is active in STEM, even if it is not framed as STEM." P: "Things that will benefit them in school." S3: "You need help like in engineering. What if you build a car? You need help? You can't do that all by yourself."

Note. UP = university personnel, CF = club facilitator, T = teacher, P = parent, S = student.

Next, Figure 1 displays a frequency chart of utterances per tiered category by stakeholder group on their conception of STEM.

Figure 1

Frequency of Utterances Per Tiered Category by Stakeholder Group on Their Conception of STEM



More utterances were captured from community stakeholders as compared to university-affiliated personnel for Category 1, with data from students comprising the highest frequency of utterances regarding their indifference towards STEM. The reverse was true for Category 2 and Category 3 in that university personnel provided more utterances than community stakeholders about STEM being a holistic discipline and its applicability and usefulness. No utterances were captured from the students or the parent for Category 2, suggesting they did not conceptualize STEM as a holistic discipline. The sections following Figure 1 provide descriptions of these categories, elaborating further as to what distinguished higher-tiered categories from the previous categories, as well as differences in conceptions between university and community stakeholders.

Category 1: Indifference Towards STEM

Fourteen utterances were assigned from the pool of meanings to this category of description. Findings from the data revealed that community stakeholders' conceptions of STEM were indifferent, as they did not know what STEM was or learning STEM was a priority. For example, two out of the three students indicated that they had never heard of STEM, while S3 communicated that she had "forgot what [STEM] stands for." The interviewed parent also articulated that she did not know the meaning of STEM, but that she was "getting a little bit from what [her son was] learning." Other stakeholders knew STEM merely as an acronym, as the interviewed teacher declared, for example, "I know what it stands for: science, technology, engineering and math and just having kids to be able to apply those four things in." Others mentioned STEM in reference to learning one of the individual disciplines (i.e., science, technology, engineering, or mathematics), rather than STEM as a whole. For instance, one of the university faculty members, UP1, asserted that "the pure definition can be in STEM, I feel, you need to be in one of the four disciplines of science, technology, engineering, and mathematics. So, like when we talk about the STEM disciplines, if you're studying biology, you're a

STEM student." The same sentiments were also echoed by two other university faculty members: UP2 and UP3. Meanwhile, one of the STEM club facilitators, CF2, articulated that "STEM is more science and math-based learning," without mentioning any regard to the technology or engineering components.

Category 2: STEM as a Holistic Discipline

Eleven utterances from the data pool were assigned to this category of description. Some stakeholders' conceptions of STEM differed from those articulated in Category 1 in that STEM was seen as an integrative and holistic discipline or a combination (i.e., multidisciplinary) of at least two of the individual STEM disciplines. All five university faculty members in this study referenced STEM as holistic or multidisciplinary. For example, UP4 stated, "when combined in the STEM fashion, there is an iteration and use of the shared relationships (among their respective knowledge, skills, or practices) that allow us to explore more convergence-based issues." Similarly, one of the community STEM club facilitators, CF1, articulated the synthesis among individual disciplines of STEM in tandem with other non-STEM related skills, stating that "STEM implementation is addressing a problem through the use of interdisciplinary skills in [each discipline of] STEM in conjunction with other cognitive and behavioral skills such as critical thinking, effective collaboration, clear and precise discourse, etc." The elementary STEM teacher also acknowledged the integration of STEM being necessary, whereas no utterances of integration or combination of disciplines were found among the interviewed parent and students of this study.

Category 3: Applicability and Usefulness of STEM

Category 3 constituted the largest amount of references made for this category of description, having a total of twenty-six utterances. The difference here between Category 2 and Category 3 in this outcome space is in regard to the applicability and usefulness of STEM. At Category 3, community stakeholders' conceptions of STEM were beyond that of merely learning the concept at the surface or definitional level. Due to nuanced variances of the twenty-six utterances within Category 3 of stakeholders' conceptions of STEM, three subcategories were formed to delineate the variation shown by stakeholder group. The three subcategories of Category 3 were related to the importance of learning STEM at home (n = 5), for school (n = 7), and their future (n = 14).

Subcategory of home-based skills. Five utterances were included in the data pool that were in reference to STEM learning that was attributed with home life. One utterance from UP3 in reference to applicability of STEM at home stated, "…taking things that they're learning in school and then figuring out how to apply them in a unique and exciting way." The remaining four utterances were from student participants on the ways they perceive STEM to be applied at home, referencing helping a parent with gardening (S2) and taking out the trash (S1) as STEM-related.

Subcategory of school-based skills. Utterances were also captured in regard to STEM learning that occurs in school. Five utterances described the learning of STEM soft skills in school that covers a wide array of disciplines including those outside of STEM (i.e., the humanities), while three utterances described the learning of STEM non-specific skills. An example of STEM soft skills, as referenced by UP5, include "working together, writing, speaking skills – in which they had to share their learning experiences with others." An example of a non-specific skill, like helping classmates in STEM, was described by S3.

Subcategory of future skills. Thirteen utterances from stakeholders related STEM applicability and usefulness in future skills. Four utterances described stakeholders' conception of STEM as a means for social advancement speaking to the inclusivity of STEM learning, as CF1 stated, "anyone and everyone is capable in learning STEM." STEM was also related to skills performed by those who worked in STEM disciplines; UP5 mentioned that "some individuals may define STEM as what I would describe as 'high-brow' STEM, referencing what scientists, engineers, and mathematicians do daily, probably in a work-setting." STEM was also related to the theme of money, with S1 making reference to an online game he played and perceived to have STEM-related content: "Then you get to make a character, and you have a lot of money." Another four utterances were coded as stakeholders' understandings of STEM as critical for success in future work. For instance, UP5 asserted that "knowledge in research and/or evaluation design is fundamental" in regard to STEM learning, while CF1 declared that "kids need to see the importance of the work that they are doing in STEM. Three utterances were in reference to the applicability and usefulness of soft skills in STEM. This subcategory involved stakeholders' conceptions of STEM as it related to soft skills that can be applied both in an out of school. For instance, S3 spoke to the importance of collaboration in STEM, and how it may be used to assist others: "how to work together, okay, and how to like help others whenever they need help." Last, commonalities in three utterances referenced other non-specific skills, focusing on the versatility of STEM for learning. For example, UP4 contended, "I would stress STEM is everywhere – and everyone is active in STEM, even if it is not framed as STEM."

Stakeholders' Conceptions of STEM Clubs

In regard to stakeholders' conceptions of STEM clubs, there were four hierarchical core categories which were labeled as follows: Category (1) *Non-STEM Related Space* with seventeen utterances; Category (2) *Promotion of STEM* with fifty utterances; Category (3) *STEM Club Learning Environment* with fifty-nine utterances; and Category (4) *Production Site of STEM Knowledge, Skills, and Enjoyment* with thirty-six utterances. Table 3 defines these core categories of description for stakeholders' conceptions of STEM Clubs, supported by examples of utterances extracted from the data across this outcome space.

Figure 2 displays a frequency chart of utterances per tiered category by stakeholder group on their conceptions of STEM clubs, highlighting differences in conceptions between university and community stakeholders. More utterances were captured from university-affiliated stakeholders than community stakeholders across all four categories. Utterances from all participants were captured in Category 1 and Category 2 regarding after school clubs being a space for non-STEM related activities and spaces that promote STEM, respectively. There were wide gaps in the frequencies of utterances between university-affiliated stakeholders and community stakeholders for Category 2 and Category 3, with university stakeholders making more utterances about the learning environment of STEM clubs. Although Category 4 still had more utterances captured from university-affiliated stakeholders, the gap among frequency differences was narrower between both groups, with student participants having about the same number of utterances as university personnel. Core categories 2, 3, and 4 were also coded into subcategories to delineate nuanced variances between utterances for data visualization. The four core categories in this outcome space are further described in the sections that follow.

Table 3

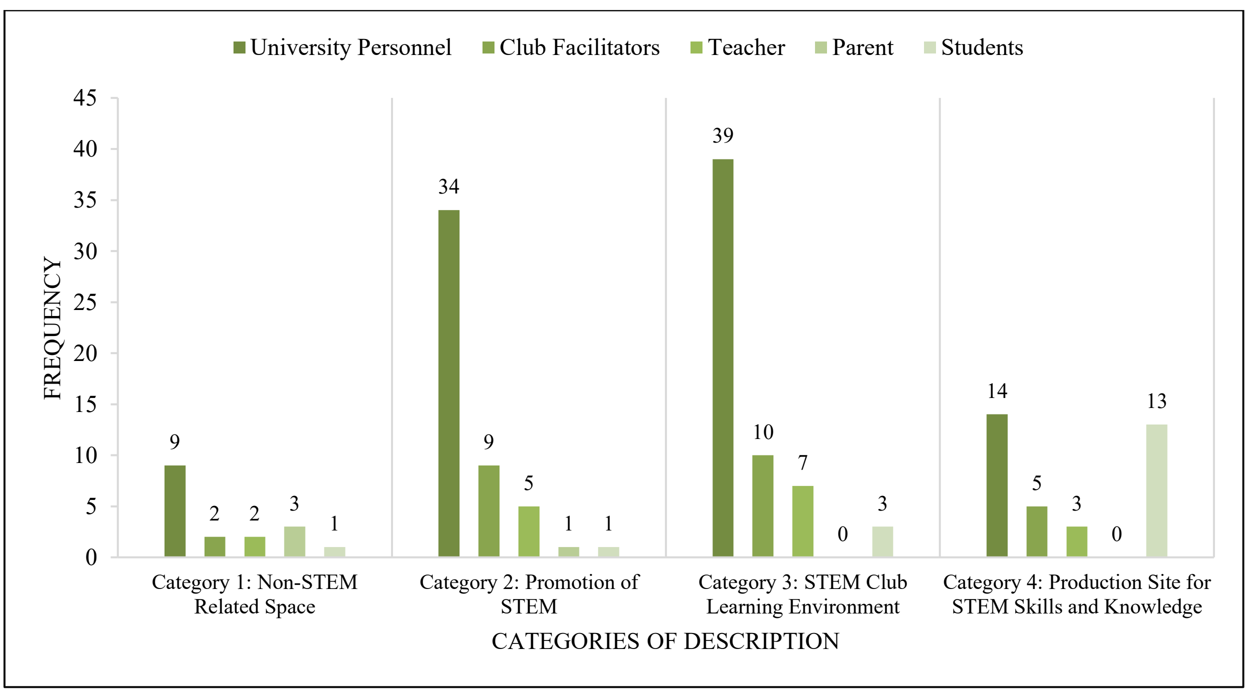
Core Categories of Description for Stakeholder's Conceptions of STEM Clubs

Core Category	Definition	Examples of supporting quotes (utterances)
Category 1: Non-STEM Related Space	Stakeholders conceptualized community STEM clubs as a safe space for students, regardless of content	P: "It's just something to do for fun." T: "It turned out to be a great thing for my higher thinkers for sure. [It] helped them emotionally because they would get frustrated."
Category 2: Promotion of STEM	Stakeholders conceptualized community STEM clubs as a space that provides students access to STEM, STEM professionals and careers, and STEM possibilities, or affect, for enhancing learning, inclusivity, and relevancy.	UP3: "The goal is to engage the average students that maybe don't have every other opportunity to engage with STEM." CF2: "I definitely noticed that the kids are excited to go to both the science and math clubs."
Category 3: STEM Club Learning Environment	Stakeholders conceptualized community STEM clubs as spaces that lie along the continuum of free choice STEM learning environments	CF1: "STEM clubs do not necessarily need to promote a static curriculum, although I do not see a problem to do so." S3: "At school like we don't have that many activities. But here it's fun because like, we do different activities for the math."
Category 4: Production Site of STEM Knowledge, Skills, and Enjoyment	Stakeholders conceptualized community STEM clubs as hubs that produce STEM knowledge, skills, and enjoyment of STEM within the club.	UP1: "You want an opportunity to expose the students to science and an opportunity to think critically, and to explore." S1: "We made tornadoes and stuff."

Note. UP = university personnel, CF = club facilitator, T = teacher, P = parent, S = student.

Figure 2

Frequency of Utterances Per Tiered Category by Stakeholder Group on Their Conception of STEM Clubs



Category 1: Non-STEM Related Space

Among the utterances assigned to Category 1, stakeholders in the CUP indicated that STEM clubs serve other purposes in addition to STEM-related content or activities. For instance, CF1 explained that in some cases, a STEM club is "a safe learning space for students who need to be kept occupied outside of normal school hours." This sentiment was also echoed in the interview with the parent participant, as she stated, "I think it's really good to have something [for the kids] to do," alluding to the fact that she was unable to pick up her son after school because she would still typically be at work at those hours. Furthermore, stakeholders indicated that the STEM club afforded an opportunity for students to establish rapport with adult mentors as role models. UP2 expressed, "If something's, you know, heavy on their mind, and it might not have anything to do with STEM, we can form those relationships with those kids."

Category 2: Promotion of STEM

Assignment of utterances for Category 2 pertained to STEM clubs as spaces that allow for the promotion of STEM. Category 2 is differentiated from Category 1 in that there were indications of STEM-related utterances to stakeholders' conceptions of STEM club, that these spaces provide a sense of direct or indirect exposure to STEM, specifically for underrepresented populations. As the teacher participant declared, STEM clubs are "exposing [kids] to things that they're not used to seeing or kind of being stretched in ways they're not used to thinking." However, due to the voluntary nature of the community STEM clubs, there are still challenges in recruiting critical learners. UP3 elaborated that "we kind of get the, you know, not the rock star students, not the lowest performing students, but somewhere kind of in between." Nonetheless, there is evidence that what is learned in STEM clubs is also being promoted at home, as the parent participant shared that her son "comes [home] to talk about what he's learned, like math, or like the science that's going on at the community center."

The utterances in Category 2 were further divided into subcategories to further capture how STEM is promoted through the club according to stakeholders' conceptions. Subcategories in Category 2 were in descending number of utterances: the inclusivity that STEM affords (n = 13); generating positive affect towards STEM through enjoyment and attitudes (n = 12), as well as cultivating interest and motivation (n = 8); gaining access to STEM professionals and careers (n = 10); engaging in STEM learning and enrichment (n = 8); garnering a relevancy of STEM (n = 4); and understanding of what is STEM (n = 2).

Category 3: STEM Club Learning Environment

Most utterances in regard to stakeholders' conceptions of STEM clubs were assigned to Category 3 which involved references to the STEM club learning environment. Utterances in Category 3 differed from Category 2 in that their attention was focused on the level of curriculum-based teaching that occurs within the STEM club. Utterances from the data revealed conflicting views within and between community members and university-affiliated personnel.

For instance, UP5 explained, "There should be a curriculum. It should build upon knowledge, skills, practices through hands-on, active experiences." However, UP4 disagreed stating, "I do not think there should be a static curriculum. This is strictly based on my definition of STEM clubs, which requires STEM club participants to be extremely fluid." Meanwhile, some stakeholders were ambivalent in regard to the decision of implementing a curriculum in a STEM club. CF2 stated, "I don't think that [STEM clubs] should necessarily be consistent because every club is catering to different kinds of people in different kinds of communities," while UP2 clarified, "The point is to have a club where we're really engaging them in those topics of STEM with any activity that we see fit for the age group."

While utterances were captured from the teacher and student participants, there were no utterances made by the parent participant that fit into this category. The utterances in Category 3 were further divided into subcategories to further capture emergent themes in relation to the STEM club learning environment according to the conceptions of stakeholders. Subcategories in Category 3 elucidated a continuum for how STEM clubs should be structured, from being completely informal (free-choice) STEM learning environments (n = 14), to being slightly structured, either guided by student-led, hands-on, enrichment (n = 13), or organized around a real-world, community-based problem (n = 11), to highly structured around a specific group or culture (n = 4), or a completely non-formal (specific curriculum) structure (n = 20).

Category 4: Production Site for STEM Skills and Knowledge

Finally, Category 4 for this outcome space was differentiated from Category 3 because this category encompassed utterances that referenced STEM skills and knowledge produced within STEM clubs, irrelevant of the STEM club curriculum, or lack thereof. Furthermore, across the stakeholder groups, there were consistencies in the understanding that STEM skills and content knowledge were learned in the STEM club. For instance, there were utterances in regard to the creation of some sort of product. The teacher participant indicated that in a previous community STEM club, students would "build cardboard and duct tape boats, that students would then get to race at the end of the week." Likewise, all students mentioned the construction of a tornado model to learn about weather, as S2 stated, "we made a tornado in a bottle." In addition to products, other soft skills and a sense of enjoyment were found to be present in STEM clubs evidenced by stakeholders' utterances. CF1 mentioned, "STEM clubs provide the opportunity to enhance global skills such as engagement, collaboration, and cooperation," alluding to soft skills acquired in STEM clubs. Meanwhile, UP3 stated that one former community STEM club student exclaimed, "And I had such a great time that I came back the next year, and now I've decided I want to be an engineer."

This category, more than any other category in this outcome space, revealed an almost equal amount of utterances between community stakeholders and university personnel. However, similar to Category 3, there were no utterances captured from the parent participant for this category. Subcategories in Category 4 were related in regard to what should be produced through STEM Clubs: knowledge (n = 16), shared enjoyment (n = 14), and skills (n = 8).

Discussion

The phenomenographical lens in the present study provided insight to the CUP stakeholders' community understandings of STEM and CUP-based STEM clubs. The analysis of the pool of meanings revealed three hierarchical categories for stakeholder conceptions of STEM: Indifference Towards STEM; STEM as a Holistic Discipline; and Applicability and Usefulness of STEM; and four hierarchical categories for stakeholder conceptions of STEM clubs: Non-STEM Related Space; Promotion of STEM; STEM Club Learning Environment; and Production Site of STEM Knowledge, Skills, and Enjoyment. There were three significant findings from the outcome space that warrant further discussion: (1) the varying degrees to which STEM is conceptualized among stakeholders; (2) stakeholders' beliefs that STEM learning is important for elementary students' futures, and afterschool community STEM clubs help promote that notion; and (3) stakeholders disagree on the learning structure of community STEM clubs.

Data suggests that there was a clear disparity between university personnel and community stakeholders' conceptions of STEM (see Figure 1). Surprisingly, university personnel in this study alluded to STEM as a more holistic discipline, contradicting what is stated in the literature that university personnel still conceptualize STEM as individual disciplines (Breiner et al., 2012). Nonetheless, Breiner and colleagues further elaborated that university personnel's conceptions of STEM are based on how STEM impacts their lives; revealing differences in this study between university STEM and STEM education faculty. That idea thus clarifies the context of this study that university personnel upheld a more integrative approach to STEM because such a conceptualization aligned with their intended learning outcomes for the STEM club; that is, informal learning via integrated STEM instruction.

As for community stakeholders, their conceptualization of STEM (Figure 1) was largely indifferent or they viewed it as generally helpful, suggesting only a rudimentary understanding of STEM. Meaning, stakeholders in the community were unaware of what STEM was in the first place. More specifically, the parent and two student participants have never heard of the concept of STEM, much less knew what the acronym stood for, reinforcing the idea that general citizens of the community do not know what STEM is (Angier, 2010).

Although the teacher participant was transparent in that she considered herself an elementary STEM teacher, she also expressed her concept of STEM merely as an acronym of the four individual disciplines. Given the increasing attention to and need for a more integrative approach to STEM (English & King, 2019), and considering that the community participants in this study were stakeholders of education at the elementary level (i.e., elementary teacher, parent of elementary student, and elementary student), there is an indication that the understanding of integrated STEM learning before the middle school years is needed. This is especially significant considering studies that have shown STEM integration being implemented at the elementary level helps to promote positive attitudes toward science in elementary students (Toma & Greca, 2018). The vast differences in a conceptualized understanding of STEM among stakeholders of the CUP, furthermore, implies that there may have been barriers and obstacles in stakeholder communication that need to be overcome (Bender, 2008). Reciprocity of understanding and substantial communication among CUP stakeholders on this particular finding would help mitigate, if not resolve, differences in conceptions of STEM. It is unsurprising then that since stakeholders hold contradicting conceptions of STEM, the same holds true regarding their conceptions of STEM clubs, as seen in the wide disparity of frequencies regarding conceptions of STEM clubs in Figure 2.

Although community stakeholders made more utterances to STEM utility at home and university personnel made more utterances to STEM utility at school, one aspect that stakeholders seemed to agree upon in regard to the conception of STEM related to the benefit of STEM for the future. The conceptualization of STEM among stakeholders was positive in that learning STEM would help promote skills that would be useful for students engaging in STEM in the long run. These soft skills include critical learning, problem solving, and collaboration—skills sought after by potential STEM employers (Prinsley & Baranyai, 2013). Moreover, stakeholders also agreed that community STEM clubs provide that space and opportunity for such STEM skills to be produced (see Figure 2), thus implying that fostering effective local community STEM clubs is beneficial not only for students, but also for the community as a whole. Indeed, as the demand for STEM skills by employers continues to increase, afterschool community STEM clubs are conducive informal learning environments to help nurture these skills in youth (Afterschool Alliance, 2015; Talafian et al., 2019).

One last significant finding from this study is that while stakeholders view community STEM clubs as significant vehicles in promoting and producing STEM learning, there were conflicting perceptions to the degree learning should be structured. Three out of five university personnel agreed with literature that afterschool STEM clubs should have a set curriculum in place (Feldman & Pirog, 2011), whereas other participants believed that STEM clubs should not be constricted to a standardized curriculum, such that learning in STEM clubs is fluid and student driven. For many of the CUP stakeholders, a sense of STEM enjoyment in a club setting was more important than the curriculum or instruction. This notion is supported in the literature that STEM clubs promote enjoyment due to hands-on experiences that are not provided in formal in-school settings (Krishnamurthi et al., 2014). However, research on effective STEM programs suggest that both the social and academic aspects of student STEM learning are positively impacted when balanced with focused objectives and curricula (NRC, 2015). In addition to consistent collaboration and communication, this research suggests that CUPs that engage in deeper conversation on the purpose of STEM clubs may establish clearer curricula and modalities for instruction to meet their elementary students' needs (knowledge, skills, and enjoyment) in STEM now and for the future.

Limitations

Phenomenography as a theoretical framework and methodological approach is often criticized for the potential bias inherent within the researcher's preconceived notions of the phenomenon (Ashworth & Lucas, 2000). Thus, to mitigate potential bias, bracketing (Tufford & Newman, 2012) was applied and accomplished through removal of stakeholder designation in the data analysis. We established credibility by building protocols and understandings from extant literature and used expert review. Inclusion of two researchers to review the data helped to mitigate introduction of bias (Trigwell, 2000; Walsh, 1994). Additional limitations included a small sample size and that participants for this study were associated with only one of the afterschool STEM clubs in the community. Sampling a larger number of stakeholders—especially teachers, parents, and students—would help determine a broader view of a community's varying conceptions of STEM. Furthermore, results examining conceptions from one afterschool community STEM club may only be significant for that geographical demographic.

Conclusion

Afterschool STEM programs have been shown to develop students' positive attitudes toward the STEM fields and interest in pursuing a career in STEM (Afterschool Alliance, 2015; Baran et al., 2019). Moreover, understanding stakeholders' conceptions of STEM is critical in establishing and sustaining the overall goals and objectives of community STEM clubs (Appel et al., 2020). The present study suggests the need for more substantive communication and collaboration among community stakeholders (i.e., university personnel, classroom teachers, community staff, parents, and students) for afterschool STEM club development and improvement (Davis et al., 2023). Evaluating and synthesizing stakeholders' conceptions of STEM would help contribute to afterschool STEM curriculum, thus strengthening the programming and implementation of STEM activities in afterschool STEM clubs. Considering stakeholder input toward afterschool STEM curriculum would enhance and inform positive outcomes in afterschool STEM clubs, thus making the clubs more attractive for student participation and parents' enrollment of their students (Salvatierra & Cabello, 2022).

The phenomenon of 'what is STEM' (Bybee 2010; Sanders 2008) and the experiences of students engaging in STEM learning within formal K-12 settings has been studied extensively (see NRC 2011; National Academy of Engineering & NRC, 2014; Shahali et al., 2016). What is less studied are the experiences of all stakeholders (students plus parents, teachers, university faculty, etc.) in informal settings such as community-based STEM clubs (Gottfried & Williams, 2013; Sahin et al., 2014). This article addressed that need through a phenomenographic examination of stakeholders' experiences in informal STEM learning via a CUP-based STEM club to explore commonalities and differences in their conceptions of STEM and experiences in informal STEM. This study found that community and university stakeholders hold varying conceptions of STEM but agree that afterschool STEM clubs are vital resources to promote STEM and enhance STEM-related life and soft skills.

However, data also disagreed in terms of the degree of curriculum structure in STEM clubs. Further studies are warranted in examining STEM clubs in various locations across various population groups to gain a clearer perspective of the STEM learning needs of a particular community.

The authors received no financial support for the research, authorship, and/or publication of this manuscript.

Richard Carlos L. Velasco (richard.velasco@ou.edu) holds a Ph.D. in curriculum and instruction from Texas Tech University and is an assistant professor of mathematics education in the department of Instructional Leadership and Academic Curriculum at the University of Oklahoma in Norman, Oklahoma, USA. His current research focuses on redressing sociopolitical issues in math education using a rehumanizing mathematics lens, investigating teacher self-efficacy for STEM education advocacy and voice in policy, and exploring curriculum in informal STEM learning spaces. Prior to higher education, Dr. Velasco was a national board certified secondary math teacher and a recipient of the Presidential Award for Excellence in Mathematics and Science Teaching.

Rebecca Hite (rebecca.hite@ttu.edu) holds a Ph.D. in science education from North Carolina State University and is an associate professor of science, technology, and engineering education in the department of Curriculum & Instruction at Texas Tech University in Lubbock, Texas, USA. Her research foci include the exploration of individual and collective science/STEM identity formation among K-12 students with a focus on under-represented, minoritized, marginalized, and under-served populations; the affordances of extended reality technologies for science and engineering teaching and learning, and the professional development of K-12 STEM master teacher leadership in policy-advocacy activities.

References

Afterschool Alliance. (2015). Full STEM ahead: Afterschool programs step up as key partners in STEM education.

Åkerlind, G. S. (2012). Variation and commonality in phenomenographic research methods. *Higher Education Research & Development*, *31*(1), 115-127. https://doi.org/10.1080/07294360500284672

Allen, P. J., Chang, R., Gorrall, B. K., Waggenspack, L., Fukuda, E., Little, T. D., & Noam, G. G. (2019). From quality to outcomes: A national study of afterschool STEM programming. *International Journal of STEM Education*, *6*(1), 37. https://doi.org/10.1186/s40594-019-0191-2

Alsop, G., & Tompsett, C. (2006). Making sense of 'pure' phenomenography in information and communication technology in education. *ALT-J Research in Learning Technology, 14*(3), 241-259. https://doi.org/10.1080/09687760600837058

Andretta, S. (2007). Phenomenography: A conceptual framework for information literacy education. *Aslib Proceedings: New Information Perspectives, 59*(2), 152-168.

Angier, N. (2010, October 4). STEM education has little to do with flowers. *The New York Times.* https://www.nytimes.com/2010/10/05/science/05angier.html

Appel, D. C., Tillinghast, R. C., Winsor, C., & Mansouri, M. (2020, August). STEM outreach: A stakeholder analysis. In 2020 IEEE Integrated STEM Education Conference (ISEC) (pp. 1-9). IEEE.

Ashworth, P., & Lucas, U. (2000). Achieving empathy and engagement: A practical approach to the design, conduct and reporting of phenomenographic research. *Studies in Higher Education*, *25*(3), 295-308.

Aslam, F., Adefila, A., & Bagiya, Y. (2018). STEM outreach activities: An approach to teachers' professional development. *Journal of Education for Teaching, 44*(1), 58-70. https://doi.org/10.1080/02607476.2018.1422618

Baran, E., Canbazoglu Bilici, S., Mesutoglu, C., & Ocak, C. (2019). The impact of an out-of-school STEM education program on students' attitudes toward STEM and STEM careers. *School Science and Mathematics, 119*(4), 223-235. https://doi.org/10.1111/ssm.12330

Bender, G. (2008). Exploring conceptual models for community engagement at higher education institutions in South Africa: Conversation. *Perspectives in Education*, *26*(1), 81-95.

Blanchard, M. R., Hoyle, K. S., & Gutierrez, K. S. (2017). How to start a STEM club. *Science Scope*, *41*(3), 88–94.

Blanchard, M. R., Gutierrez, K. S., Hoyle, K. S., Painter, J. L., & Ragan, N. S. (2017). *Rural, underrepresented students' motivation, achievement, and perceptions in afterschool STEM clubs.* Paper presented at the meeting of the European Science Education Research Association, Dublin, Ireland.

Blank, R. K. (2012). *What is the impact of decline in science instructional time in elementary school.* Noyce Foundation.

Breiner, J. M., Harkness, S. S., Johnson, C. C., & Koehler, C. M. (2012). What is STEM? A discussion about conceptions of STEM in education and partnerships. *School Science and Mathematics*, *112*(1), 3-11. https://doi.org/10.1111/j.1949-8594.2011.00109.x

Bybee, R. W. (2010). What is STEM Education? *Science, 329*(5995), 996.

Bybee, R. W., & Fuchs, B. (2006). Preparing the 21st century workforce: A new reform in science and technology education. *Journal of Research in Science Teaching, 43*(4), 349-352. https://doi.org/10.1002/tea.20147

Ching, Y. H., Yang, D., Wang, S., Baek, Y., Swanson, S., & Chittoori, B. (2019). Elementary school student development of STEM attitudes and perceived learning in a STEM integrated robotics curriculum. *TechTrends*, *63*(5), 590-601. https://doi.org/10.1007/s11528-019-00388-0

Cope, C. (2004). Ensuring validity and reliability in phenomenographic research using the analytical framework of a structure of awareness. *Qualitative Research Journal*, *4*(2), 5-18.

Curwood, S. E., Munger, F., Mitchell, T., Mackeigan, M., & Farrar, A. (2011). Building effective community-university partnerships: Are universities truly ready? *Michigan Journal of Community Service Learning*, *17*(2), 15-26.

Davis, K., Fitzgerald, A., Power, M., Leach, T., Martin, N., Piper, S., Singh, R., & Dunlop, S. (2023). Understanding the conditions informing successful STEM clubs: What does the evidence base tell us?. *Studies in Science Education*, *59*(1), 1-23.

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

Duodu, E., Noble, J., Yusuf, Y., Garay, C., & Bean, C. (2017). Understanding the delivery of a Canadian-based after-school STEM program: A case study. *International Journal of STEM Education*, *4*(1), 1-11. https://doi.org/10.1186/s40594-017-0083-2

English, L.D., & King, D. (2019). STEM Integration in sixth grade: Designing and constructing paper bridges. *International Journal of Science and Mathematics Education, 17,* 863–884. https://doi.org/10.1007/s10763-018-9912-0

Eshach, H. (2007). Bridging in-school and out-of-school learning: Formal, non-formal, and informal education. *Journal of Science Education and Technology*, *16*(2), 171-190. https://doi.org/10.1007/s10956-006-9027-1

Feldman, A., & Pirog, K. (2011). Authentic science research in elementary school after-school science clubs. *Journal of Science Education and Technology*, *20*(5), 494-507. https://doi.org/10.1007/s10956-011-9305-4

Foster, K. M., Bergin, K. B., McKenna, A. F., Millard, D. L., Perez, L. C., Prival, J. T., Rainey, D. Y., Sevian, H. M., VanderPutten, E. A., & Hamos, J. E. (2010). Partnerships for STEM education. *Science Education*, *329*(5994), 906-907.

Gandhi-Lee, E., Skaza, H., Marti, E., Schrader, P. G., & Orgill, M. (2017). Faculty perceptions of student recruitment and retention in STEM fields. *European Journal of STEM Education*, *2*(1), 2-11. http://dx.doi.org/10.20897/esteme.201702

Gottfried, M. A., & Williams, D. (2013). STEM club participation and STEM schooling outcomes. *Education Policy Analysis Archives*, *21*(79), 1-27. http://epaa.asu.edu/ojs/article/view/136

Hernandez, D., Rana, S., Alemdar, M., Rao, A., & Usselman, M. (2016). Latino parents' educational values and STEM beliefs. *Journal for Multicultural Education, 10*(3), 354-367. https://doi.org/10.4018/978-1-6684-7771-7.ch009

Hite, R., Midobuche, E., Benavides, A. H., & Dwyer, J. (2018). Third space theory: A theoretical model for designing informal STEM experiences for rural Latina youth. In T. T. Yuen, E. Bonner, & M. G. Arreguin-Anderson (Eds.), (*Under)Represented Latin@s in STEM: Increasing Participation Throughout Education and the Workplace* (pp. 189-202). Peter Lang Publishing.

Hite, R., & White, J. (2019). Balancing profits and conservation: A human environmental impact PBL for upper elementary and middle grades STEM club students. *Science Activities, 56*(3), 88-107. https://doi.org/10.1080/00368121.2019.1693950https://doi.org/10.1080/1533015X.2021.1986431https://journals.uc.edu/index.php/cye/article/view/6155

Krishnamurthi, A., Ballard, M., & Noam, G. G. (2014). Examining the impact of afterschool STEM programs. Noyce Foundation. https://files.eric.ed.gov/fulltext/ED546628.pdf

Kvale, S. (1996). *InterViews: An introduction to qualitative research interviewing.* Sage Publications.

Lavy, V. (2010). *Do differences in school's instruction time explain international achievement gaps in math, science, and reading?: Evidence from developed and developing countries*. National Bureau of Economic Research.

Limburg, L. B. (2008). Phenomenography. In L. M. Given (Ed.), *The SAGE encyclopedia of qualitative research methods* (pp. 611-614). Sage. https://doi.org/10.4135/9781412963909.

Marton, F. (1981). Phenomenography—describing conceptions of the world around us. *Instructional Science*, *10*(2), 177-200. https://doi.org/10.1007/BF00132516

Marton, F. (1986). Phenomenography—a research approach to investigating different understandings of reality. *Journal of Thought, 21*(3), 28-49.

Marton, F., & Booth, S. (1997). *Learning and awareness.* Lawrence Erlbaum.

Moll, L. C., Amanti, C., Neff, D., & Gonzalez, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory into practice*, *31*(2), 132-141. https://doi.org/10.1080/00405849209543534

Mullet, D. R., Kettler, T., & Sabatini, A. (2018). Gifted students' conceptions of their high school STEM education. *Journal for the Education of the Gifted*, *41*(1), 60-92. https://doi.org/10.1177/0162353217745156

National Academy of Engineering, & National Research Council. (2014). *STEM integration in K-12 education: Status, prospects, and an agenda for research.* The National Academies Press.

National Research Council. (2014). *Successful K-12 STEM education: Identifying effective approaches in science, technology, engineering, and mathematics*. National Academies Press.

National Research Council. (2015). *Identifying and supporting productive STEM programs in out-of-school settings*. National Academies Press.

NGSS Lead States. 2013. *Next Generation Science Standards: For States, By State*s. National Academies Press.

Ornek, F. (2008). An overview of a theoretical framework of phenomenography in qualitative education research: An example from physics education research. *Asia-Pacific Forum on Science Learning and Teaching, 9*(2), 1-14.

Otter. (2020). *Otter is where conversations live.* [Computer software] https://otter.ai/

Plasman, J. S., Gottfried, M., & Williams, D. (2021). Following in their footsteps: The relationship between parent STEM occupation and student STEM coursetaking in high school. *Journal for STEM Education Research, 4*(1), 27-46. https://doi.org/10.1007/s41979-020-00040-0

https://doi.org/10.4018/978-1-6684-7771-7.ch005https://ejrsme.icrsme.com/article/view/20429Prinsley, R. T., & Baranyai, K. (2013). STEM skills in the workforce: What do employers want? *Office of the Chief Scientist: Occasional Paper Series,* 9, 1-4.

Radloff, J., & Guzey, S. (2016). Investigating preservice STEM teacher conceptions of STEM education. *Journal of Science Education and Technology*, *25*(5), 759-774. https://doi.org/10.1007/s10956-016-9633-5

Ring, E. A., Dare, E. A., Crotty, E. A., & Roehrig, G. H. (2017). The evolution of teacher conceptions of STEM education throughout an intensive professional development experience. *Journal of Science Teacher Education*, *28*(5), 444-467. https://doi.org/10.1080/1046560X.2017.1356671

Sahin, A., Ayar, M. C., & Adiguzel, T. (2014). STEM related after-school program activities and associated outcomes on student learning. *Educational Sciences: Theory and Practice*, *14*(1), 309-322.

Salvatierra, L., & Cabello, V. M. (2022). Starting at Home: What does the literature indicate about parental involvement in early childhood STEM education? *Education Sciences, 12*(3), 218.

Sanders, M. (2008). STEM, STEM education, STEMmania. *The Technology Teacher, 68*(4)*,* 20-26.

Shahali, E. H. M., Halim, L., Rasul, M. S., Osman, K., & Zulkifeli, M. A. (2016). STEM learning through engineering design: Impact on middle secondary students' interest towards STEM. *EURASIA Journal of Mathematics, Science and Technology Education*, *13*(5), 1189-1211. https://doi.org/10.12973/eurasia.2017.00667a

Talafian, H., Moy, M. K., Woodard, M. A., & Foster, A. N. (2019). STEM identity exploration through an immersive learning environment. *Journal for STEM Education Research*, *2*(2), 105-127. https://doi.org/10.1007/s41979-019-00018-7

Tay, J., Salazar, A., & Lee, H. (2018). Parental perceptions of STEM enrichment for young children. *Journal for the Education of the Gifted*, *41*(1), 5-23. https://doi.org/10.1177/0162353217745159

Toma, R. B., & Greca, I. M. (2018). The effect of integrative STEM instruction on elementary students' attitudes toward science. *Eurasia Journal of Mathematics, Science and Technology Education*, *14*(4), 1383-1395. https://doi.org/10.29333/ejmste/83676

Trigwell, K. (2000). A phenomenographic interview on phenomenography. In J. Bowden & E. Walsh (Eds.), *Phenomenography* (pp. 62-82). RMIT University Press.

Tufford, L., & Newman, P. (2012). Bracketing in qualitative research. *Qualitative Social Work*, *11*(1), 80-96. https://doi.org/10.1177/1473325010368316

Walsh, E. (1994). Phenomenographic analysis of interview transcripts. In. J. Bowden & E. Walsh (Eds.), *Phenomenographic research: Variations in method* (pp. 17-30). RMIT University Press.

Appendix A

Questions for University Personnel	Questions for Teacher	Questions for Parent	Questions for Students
1. How do you define 'STEM?'	1. How do you define 'STEM?'	1. How do you define 'STEM?'	1. What do you think of when you hear 'STEM?'
2. How do you define a 'STEM club?'	2. How do you define a 'STEM club?'	2. How do you define a 'STEM club?'	2. What do you think of when you hear 'STEM club?'
3. What do you expect students to learn in STEM clubs?	3. What do you expect your students to learn in STEM clubs?	3. What do you expect your child(ren) to learn in STEM clubs?	3. What do you expect to learn when you go to STEM club?

Appendix B

Preliminary Categories of Description (n=17) in Alphabetical Order

Category
Acknowledgment of different definitions of STEM
Affective learning in STEM
Characteristics of STEM club
Contents of STEM club
Does not know STEM
Learner-centered STEM club
Prior experiences in STEM
Purpose of community STEM club
STEM clubs are thematic
STEM as individual disciplines
STEM as interrelated disciplines
STEM club learning environment
STEM exposure
STEM is a part of everyday life
STEM is professional work-setting related
Teaching-centered STEM club
Unfocused STEM clubs

ELECTRONIC JOURNAL FOR RESEARCH
IN SCIENCE & MATHEMATICS EDUCATION
VOL. 27, NO. 3, 96-121

Integrating Environmental Knowledge into a Short Interdisciplinary Course on Sustainability

Pankaj Sharma
Purdue University

Eric D. Deemer
Purdue University

Jane Lu Hsu
National Chung Hsing University

ABSTRACT

This study examines the efficacy of a novel college sustainability course in promoting relevant environmental knowledge and interest in careers related to environmental aspects within the Taiwanese educational context. The core content of the course covers the essential concepts of sustainability and introduces students to environmental issues and their interrelation with the nexus of food, energy, and water, as well as related economic and social issues. This action competence-focused course was designed to allow students to develop their understanding of sustainability through a combination of engaging lectures, novel group activities, case studies, exercises, and team projects. The sample consisted of 44 Taiwanese undergraduate and graduate students majoring in the science, technology, engineering, and mathematics (STEM) disciplines. Participants' interest in STEM careers and perceived knowledge were measured by a pre-test and post-test administered before and after the program, respectively. Analyses of variance, correlation analyses, and cross-lagged panel regression analyses were conducted to test four hypotheses. Results of repeated-measure analysis of covariance indicated that knowledge increased significantly from pre-test to post-test, but not career interest. Results of a cross-lagged panel regression analysis also indicated that pre-test knowledge was a significant positive predictor of post-test career interest. By creating an engaging class atmosphere and promoting experiential self-learning activities, this course was highly effective in enhancing students' knowledge of key sustainability aspects. Implications for interest development theory and sustainability pedagogy are also discussed.

Keywords: environmental knowledge, sustainability, sustainable development goals, interdisciplinary course, STEM, career planning, Taiwan

Introduction

Sustainable development refers to meeting the needs of the present without compromising the ability of future generations to satisfy their own needs (Brundtland, 1987; Borowy, 2021). Eight millennium development goals were adopted for a 15-year period (United Nations Report, 2015), and these goals have been broadened into 17 sustainable development goals (United Nations Report, 2016). The United Nations declared 2004-2015 the Decade for Sustainable Development and introduced the Principles for Responsible Management Education to enhance and extend

Sustainability into mainstream education. Education for sustainability allows for the development of knowledge, skills, values, and broader perspectives necessary for people to act in ways that contribute to sustainable living by better understanding of environmental, social, cultural and economic systems and their interdependence. Evans et al. (2017) emphasized the need for a more systematic and cross-disciplinary approach to sustainability education with interactive teaching and learning methods. Aikens et al. (2016) noted that not all sustainability programs focus on topics such as climate change policy and intersectionality. Numerous sustainability courses lack interdisciplinary coverage on key topics. The learning processes include collaboration and dialogue; engaging the "whole system"; innovating curriculum as well as teaching and learning experiences; and processes of active and participatory learning, with assessment of learning processes and outcomes relationship.

Programs that focus on sustainable development are structured to achieve the following goal for students: encourage students to enhance their environmental knowledge to gain a broad understanding of sustainable development goals. Environmental educators have four responsibilities while avoiding indoctrination: (1) help learners understand why sustainable development ought to be of interest to them; (2) help students gain multiple perspectives on issues; (3) help students understand what they are learning and its significance; and finally (4) encourage students to continue to think about what to do, individually and socially, to keep their own and others' options open (Qablan et al., 2011; Scott, 2002). Kasimov et al. (2002) explained that guiding principles related to sustainability, environmentalism, economics, and social well-being need to be emphasized in education for sustainability. Owusu et al. (2017) stated that business students' interest in environmental issues is positively related to their environmental literacy levels; moreover, once they have a deeper understanding of environmental issues, they are more likely to be involved in environmental activities.

Charatsari and Lioutas (2018) found that participation in a short environmental education course provides students with higher levels of environmental knowledge and a more holistic understanding of the environment, while it also increases the impacts of environmental education in higher education and recommends the introduction of environmental education in curricula to facilitate the development of environmental thinking. A novel concept has been implemented by creating the University Regional Research Consortium for environmental monitoring and protection to improve education applicability (Şterbuleac & Toma, 2019). By forming relations with community members and decision-makers, a university offers ways of managing current environmental challenges (Şterbuleac & Toma, 2020). Choudhary et al. (2019) report that students with a science background are more likely to have higher levels of interest, knowledge, participation, and contribution toward the environment compared to students with non-science backgrounds.

Jensen and Schnack (1997) recommended the concept of *action competence* in environmental education, making the argument that environmental issues are deeply rooted in societies. Action competence was defined as relevant *knowledge*, *self-efficacy*, and *willingness* in three constructs: (1) knowledge of action possibilities; (2) confidence in one's own influence; and (3) the willingness to act (Breiting & Mogensen,1999; Jensen & Schnack, 1997; Olsson et al., 2020). Olsson et al. (2020) further defined action competence:

> In the action competence concept, which we here define as a latent capacity among individuals, the need for meaningful actions is described as the willingness to act for sustainability (p. 745).

Sass et al. (2021) emphasized the importance in action competence in sustainable development (ACiSD) and considered it as the desired outcome of education for sustainable development (ESD). Environmental education enables students to act on both societal and personal levels. Furthermore, Miller et al. (2021) found that, after an internship sustainability program, university students' environmental knowledge increases.

Taiwan, in particular, faces significant sustainability challenges. Due to its comparative disadvantage in agricultural production, the self-sufficiency rate in calorie-based calculations of staple food was 26.31% in 2021 (Taiwan Ministry of Agriculture, 2023). Another major challenge is related to the environment. As one of the most densely populated places in the world, Taiwan has to deal with the growing problem of microplastics, and food and electronic waste. In addition, its renewable energy sufficiency is also a concern as the country's electricity generation still mainly relies on thermal energy, while only less than 10% is from renewable sources (Bureau of Energy, Ministry of Economic Affairs, 2022). By 2025, the planned target is to have 20% renewable energy. In light of the nuclear disaster that occurred in March 2011 in Fukushima, Japan, Taiwan initiated the decommission of its nuclear power plants, which contributes to less than 10% of the nation's electricity. Plans are underway to replace these old nuclear power plants with thermal, hydropower, and wind energy.

Moreover, research has demonstrated that while students in Taiwan frequently show interest in sustainability and environmental issues, they rarely connect it to the nexus of food, water, and energy or to government policy, focusing instead on individual actions, such as recycling (Hsu & Pivec, 2021). They also appear to lack sufficient knowledge of climate change (Li & Liu, 2021). However, individual actions and behavior alone would be insufficient to solve the issues relating to sustainability and the linkages between energy, water, and food. The gravity of the environmental situation in Taiwan requires the development of a workforce with comprehensive sustainability knowledge and the occupational skills needed to address these crucial challenges. To do this, educators are encouraged to stimulate students' interest in pursuing careers in the fields of science, technology, engineering, and mathematics (STEM), in which environmental innovations will primarily be developed (Vennix et al., 2018). Furthermore, sustainable development courses should emphasize the linkages between food, water, and energy, and their relationship to government policy to ensure that sustainability knowledge and interests are broad enough to enable Taiwanese students to comprehensively understand the environmental issues that face Taiwan and the world beyond it.

The purpose of this study was to examine the efficacy of a novel college sustainability course in promoting relevant knowledge and STEM career interests related to environmental aspects of sustainability within a Taiwanese educational context. In Taiwan, short integrated and interdisciplinary courses focusing on sustainability is a fairly new concept in local education. Moreover, our course was taught in English and it delivers the cutting-edge knowledge in sustainability to students with its original formality in content. Hence, our course was considered as a novel college sustainability course since its concept and teaching method were new to college students in Taiwan. The four-day course was titled *Sustainable Development: Environmental, Economical, Managerial, and Health Perspectives: Inspiring Future Leaders in Sustainable Development,* and course schedules are described in Appendix I. The primary objectives of the course were to improve students' knowledge of the following:

(1) The major challenges of energy, food, and water consumption
(2) The management of natural resources and sustainability
(3) The impact of technology on sustainability and climate change
(4) Economic, business, and managerial perspectives on sustainable development.

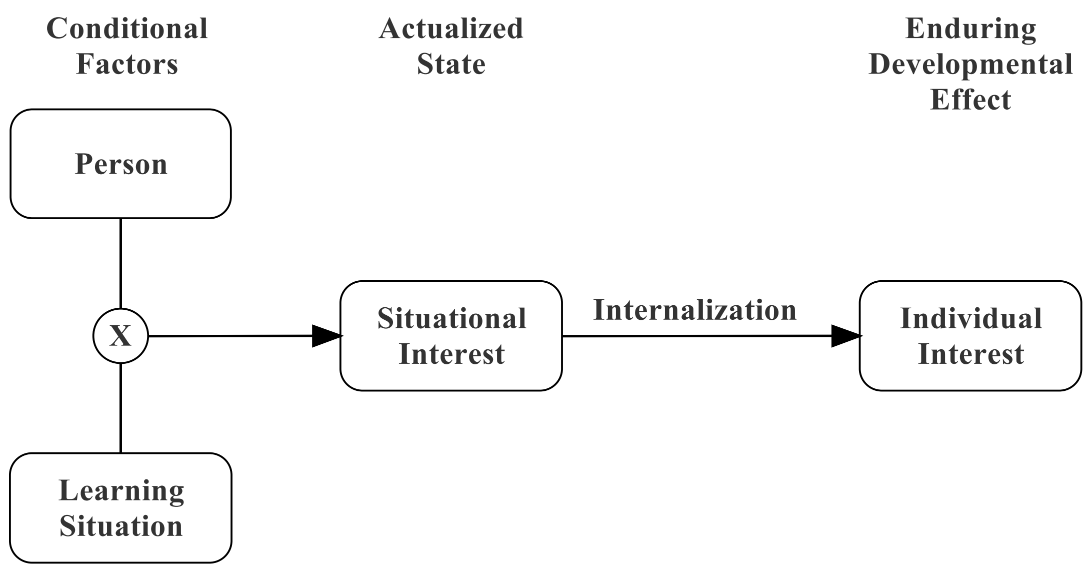
STEM education has a fundamental role in advancing technology, medicine, sustainability, agriculture, national security, economy, and society. STEM courses encourage finding the most viable global problems without necessarily considering the most sustainable option. To prepare STEM students to become potential change agents, they must learn to apply their knowledge to the three pillars of sustainability: economic, environmental, and social. An integration of sustainability into STEM courses is much needed (Zizka et al., 2021). We note that students may enter STEM careers not only to become practitioners but also to educate others about sustainability. However, in this study, we limited our focus to students who were primarily interested in careers in STEM practice rather than STEM pedagogy. This article has been organized into the following sections: theoretical framework, teaching sustainability, hypothesis development, teaching methodology, results, discussion, limitations, future research, and conclusions.

Theoretical Framework

We integrated two theoretical perspectives to examine the sustainability course: the Person–Object Theory of Interest (POI) (Krapp, 2002) and Alexander's Model of Domain Learning (MDL) (Alexander, 2004; Alexander et al., 1994). The POI posits that interest develops as a function of the relationship between a person and their life space. See Figure 1 for the person-object theory of interest model.

Figure 1

Model of the Person–object Theory of Interest Development



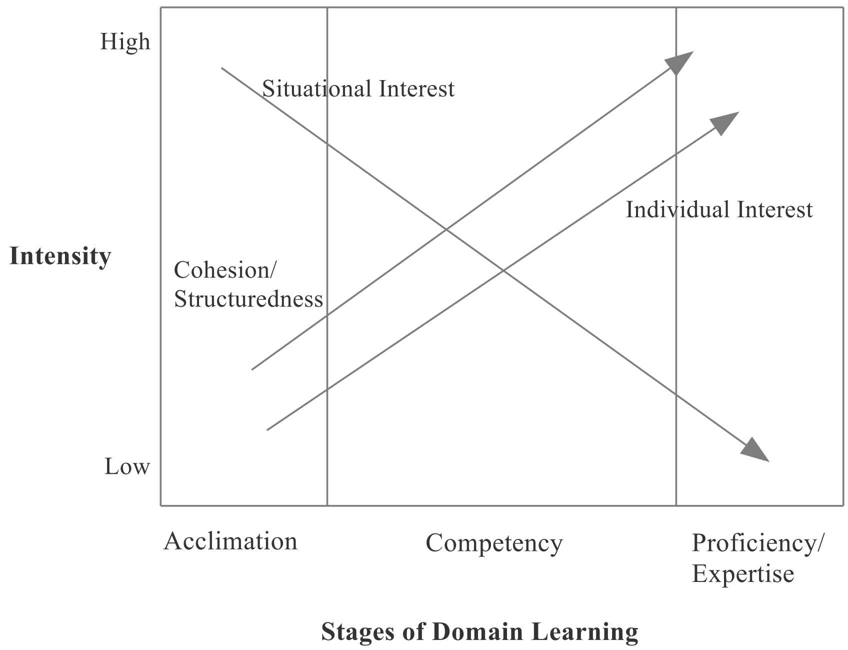
Note. Situational interests in this study are defined as the interest shown by students for the various problem-based learning emphasized in the course. Figure adapted from "Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective," by A. Krapp, 2002, p. 398. Copyright 2002 by Elsevier

From a personality theory perspective, this interest is assumed to develop through the connections made between an object of interest and the attitudes and feelings that define an individual, which are stable across situations and time. This type of interest is referred to as *individual interest.* However, interest can also be activated through transient interaction with an object or task. This structural component of interest, referred to as *situational interest*, is primarily determined by the features of the task and is characterized by heightened attention and affective engagement (Hidi, 2000; Hidi & Harackiewicz, 2001). The interaction of the person with the object—which can represent a concrete object, subject matter, or idea (Krapp & Prenzel, 2001)—is theorized to lead to the development of situational interest, which is then internalized over time as individual interest.

The MDL builds upon the POI by incorporating both situational and individual interests into a model that posits both as a function of domain and topic knowledge. See Figure 2 for information.

Figure 2

Model of Domain Learning (Alexander, 2004, pp. 273–298)



Note. From "A model of domain learning: Reinterpreting expertise as a multidimensional, multistage process." In D. Y. Dai and R. J. Sternberg (Eds.), Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development (pp. 273–298). Lawrence Erlbaum. Copyright by Lawrence Erlbaum.

According to this theory, domain learning occurs across three stages: acclimation, competency, and proficiency or expertise. Learners who are acclimating to the knowledge domain are theorized to possess high situational interest due to the novelty of the task. As learners develop competence in a domain, their situational interest declines. However, with declining situational interest, there is a concomitant increase in individual interest as the learner's competence increases and learning becomes more formalized and structured. We grounded the current study in this theoretical framework because other theories of interest development (e.g., Lent et al., 1994) focus exclusively on trait-like interest development rather than focusing equally on trait- and situation-based determinants of individual interest.

Empirical evidence supports both theoretical propositions. For instance, Tröbst et al. (2016) indicated that elementary students' perceptions of science instruction by their teachers were positively associated with the development of individual interest. Moreover, Tröbst et al. (2016) demonstrated that situational interest mediated this relationship, thus indicating that situational interest functions as a temporal antecedent of individual interest. Research has also shown that interaction with particular mathematical tasks (e.g., calculation) is positively associated with the development of interest in these tasks, as well as mathematics interest in general (Ufer et al., 2017). Studies employing pretest-posttest designs among college-aged students also suggest that effective teaching strategies can increase situational interest in STEM domains such as marine biology (e.g., Seidelin et al., 2021). Research focusing specifically on the MDL has also supported its theoretical propositions. In a sample of college students, Murphy and Alexander (2002) found that pretest interest in psychological subject matter was associated with increased interactive knowledge, which in turn predicted an increase in posttest subject-matter knowledge. Studies have also supported the theory's tenets among undergraduate students in the life and physical sciences. For instance, Alexander et al. (1995) found that participants with greater knowledge of immunology displayed higher interest in and recall of information on the topic than participants with less immunology knowledge. Similarly, Alexander et al. (1994) demonstrated that domain knowledge of physics was a positive predictor of both interest in physics and later recall of physics information. Despite these supportive findings, there remains a dearth of research on these interest development models among undergraduate students in STEM, which underscores the critical need for the present research. Moreover, there is not only a need to increase students' interest in conducting research in these STEM domains. There is also a need to raise students' awareness of the connections of STEM research topics to broader non-STEM issues (e.g., public policy, economics, politics) that are also implicated in discussions of environmental sustainability. The current research aims to unify these topics in this way.

Current Study

Numerous studies have supported the efficacy of interventions aimed at altering both environmental sustainability (e.g., Lehman & Geller, 2004) and science-related attitudes (e.g., Deemer & Sharma, 2019). Therefore, our first two hypotheses are as follows:

Hypothesis 1: Participants' perceptions of knowledge of environmental and sustainability issues will increase significantly from pre-test to post-test.

Hypothesis 2: Participants' individual interest in STEM careers will increase significantly from pre-test to post-test.

Given that prior knowledge provides a foundation for the development of individual interest (Alexander, 2004)—a finding that has been demonstrated consistently across STEM subject domains (e.g., Durik & Matarazzo, 2009; Tapola et al., 2013)—we expected that antecedent perceptions of knowledge would be associated with subsequent career interest. Hence,

Hypothesis 3: Participants' perceptions of knowledge of environmental sustainability at pre-test will be a significant positive predictor of STEM career interest at post-test.

Finally, MDL suggests that the relationship between knowledge and interest should strengthen with increasing exposure to an academic domain (Alexander, 2004). Lawless and Kulikowich (2006) obtained support for this assertion by finding that the correlation between these variables increases as students advance from undergraduate to graduate study. Therefore,

Hypothesis 4: The positive correlation between post-test knowledge and post-test interest will be significantly stronger than the positive correlation between pre-test knowledge and post-test interest.

Method

Course Contexts

The data presented here was collected during a four-day course that met during fall break in September 2018. Prior to the program, participants completed a pre-test which consisted of demographic, STEM career interests, and perceived knowledge questionnaires. In terms of the curriculum, the sustainability module was presented on the first day, the energy and technology modules were presented on the second day, the managerial, global health, and nexus modules were delivered on the third day, and students presented the results of their group projects on the fourth day. At the conclusion of all program activities, the participants were asked to complete a post-test survey that assessed their STEM career interests and perceived knowledge. We administered all surveys online.

We implemented revised course contents and teaching methodology from the pilot study conducted in 2016 through 2017. A more comprehensive questionnaire was used in administrating the survey since the course design covers a wide range of topics relevant to STEM careers, which may be of interest in career development for participants. The sustainability short course lasted for four days, which is not a typical duration for college-level courses in Taiwan. A typical university course is of three credits, with an approximately three-hour weekly class across 15-18 weeks. A short learning course is offered with less lecturing hours. Effectiveness of short learning courses are evaluated for the benefit of improving the students' self-efficacy (Judge et al., 2020); improving knowledge and skills (Argimon-Pallàs et al., 2011); and improving student communication skills (Hazelton et al., 2009). In this research, our course was considered short as it was offered for one credit hour, which is 18 hours in total. The short course provided enough data to test four hypotheses, all of which were grounded in the POI and MDL frameworks.

Our teaching methodology was developed in a short course for university students and implemented over the course of 3 years (i.e., 2016, 2017, and 2018) at two national universities in Taiwan. Revised course data was collected in 2016 and 2017 to modify and refine the course content and teaching methodology. Only data collected in 2018 are presented in this research. Despite a small sample size, we subjected the data collected in 2018 to empirical investigation to assess the preliminary efficacy of the pedagogical intervention.

Participants

The sample consisted of a total of 44 Taiwanese students majoring in STEM disciplines: 30 undergraduate students (five freshmen, six sophomores, nine juniors, and 10 seniors) and 14 graduate students. The majority of students (n = 37) were from National Chung Hsing University (NCHU) in Taiwan because few students were recommended to join the course by their advisors at other universities. As long as their majors were within the spectrum of STEM disciplines and they had language proficiency in English, they were qualified to participate in the program. Twenty-six participants identified as female and 18 identified as male. The participants' mean age was 20.49 years, with an average of 2.07 college courses (SD = 1.02), in which they had previously learned about some aspects of energy, water, food, environment, and health-related topics. Most of these courses were in economics (n = 25), followed by biology (n = 20), earth science (n = 13), chemistry (n = 12), technology (n = 9), physics (n = 3), and sociology (n = 1).

Measures

STEM Career Interest

Participants' interest in STEM careers was measured with a pre-test and post-test administered before and after the program using five items developed by the first author. These items asked participants to rate their interest in the following career domains: (a) science, (b) technology, (c) engineering, (d) environment, and (e) mathematics. Specifically, the item for each domain was as follows: "How would you rate your interest in (science, technology, engineering, environment, or mathematics)?" Items were rated on a Likert-type scale ranging from 1 to 4 (1 = never considered; 2 = very weak; 3 = moderate; 4 = very strong) (e.g., see Appendix II and III) and summed to create a composite score. Cronbach's alpha coefficients of .70 and .88 were obtained at pre-test and post-test, respectively.

Perceived Knowledge

Participants' perceptions of their knowledge of various issues related to environmental education were assessed using 10 items developed by the first author (see Appendix II). Items were rated on a Likert-type scale ranging from 1 to 5 (1 = not much; 2 = very little; 3 = a moderate amount; 4 = a considerable amount; 5 = a great deal). The scale exhibited excellent internal consistency reliability at pre-test (α = .86) and post-test (α = .93).

Course Curriculum

A pedagogical strategy is to focus on action competence in delivering courses. Education for sustainability is about engaging students with the world they live and developing the ability to act for a sustainable future. The teaching and learning approaches can include experiential learning, cooperative learning, problem-based learning, and inquiry-based learning. These approaches prepare students to be active participants, empower their capability of deliberating causes and effects, and constructs their visions for finding strategies toward solving sustainably problems. There has been a sharp increase in interdisciplinary teaching of sustainable development within environmental education (Chen & Liu, 2020; Jensen & Schnack, 1997). Examples of topics included in courses are sustainability or sustainable development, natural resources, environmental health, and global warming. Most of the courses were at the intervals of one week to one semester and reported positive outcomes (Chen & Lu, 2020). Pedagogical strategies and active learning included role-plays and simulations, group discussions, stimulus activities, debates, critical incidents, case studies, reflective accounts, critical reading and writing, problem-based learning, fieldwork and outdoor learning, modeling good practices, and seeing the big picture (Tilbury, 2011).

The lectures were chosen to provide a broader understanding of sustainability and sustainable development goals. A four-day framework was chosen to fit the course during fall break. We organized the lectures into two sessions per day, and each session focused on a different aspect of sustainability. At the start of the course, an "ice-breaker" session facilitated team-building. Appendix I shows details of the course curriculum and activities (introduction/icebreaker, lectures, open discussion, break, group projects, team buildings games, case studies, group presentations, and prize distribution). Interactive learning was incorporated as instructors encouraged students to ask questions and work in teams to respond to questions. Peer-to-peer student interactions were encouraged in providing broader perspectives and applying sustainability/sustainable development knowledge to increase their STEM career interests.

The course was divided into a series of lectures or sessions that discussed the various aspects of sustainability and its relationship with the environment, the energy, food and water nexus, global health, and related economic policy. Our lectures also included several pertinent case studies. The various topics of discussion or sessions were as follows.

Global Sustainability Challenges of the 21st Century. We defined sustainable development (Redclift, 2006) and discussed its impacts on policies relating to energy, climate and environment, water, and food security.

Energy Resources: Fossils and Renewables. In the first part of the lecture, students were introduced to three primary types of energy sources—fossil, renewable, and nuclear. The second part of the lecture included renewable forms of energy (i.e., solar, hydropower, wind, geothermal, and biomass).

Electricity Grid and Energy Storage. Two core concepts were covered in this lecture: energy storage (Luo et al., 2015) and electrical grids (Diamantoulakis et al., 2015).

Energy Utilization. We introduced students to the main energy-consuming sectors (U.S. Energy Information Administration, 2018): the industrial sector, the transportation sector, the residential sector, the commercial sector, and the electric power sector.

Food and Water. We introduced the concept of food security (Misra, 2014). We then outlined the four main dimensions of food security: availability, access, utilization, and stability. We also discussed the sources and usage of water, the water cycle, water availability, and the water footprint network (Gunders, 2012; Provide Access to Clean Water, 2018).

Environment and Climate. This session introduced students to the global environmental challenges brought about by human activity, the growth of the human population, and the increased stress that humans place on the biosphere (National Research Council, 2001; Rockström et al., 2009).

Economics and Business. In this session, students studied energy economics and commodities. In addition, we discussed policies, regulations, and the concept of a circular economy (McKinsey Center for Business and Environment, 2016; Castillo et al., 2018; Coyle & Simons, 2014).

Global Health. In this lecture, health as it refers to the global population, was discussed. Global health is defined as the "area of study, research, and practice that places priority on improving equity in health for all people worldwide" (Kaplan et al., 2009, para. 1).

NEXUS Perspective: Interdependencies and Interconnectedness. The water, energy, and food security nexus refers to the inextricable link between water security, energy security, and food security—actions in any one area usually have impacts in one or both of the others (D'Odorico et al., 2018). As the world population approaches nine billion and demands for basic services and the desire for higher living standards increase, the need for more conscious stewardship of the vital resources required to meet those demands and desires has become both more obvious and urgent. A nexus approach can enhance water, energy, and food security by increasing efficiency, reducing trade-offs, building synergies, and improving governance across sectors. A nexus approach is a requirement for effective policies to benefit society.

Case Studies on Environmental Impact. Several case studies were presented throughout the course. The selected case studies investigated the effects of industrial development activities or disasters on the environment, communities, and public health. These included the Fukushima disaster (American Nuclear Society, 2012), the BP Oil spill (National Academy of Engineering & National Research Council, 2012), and Nam Theun 2, a hydropower dam project in Laos (The World Bank, 2018).

Teaching Methodology

In the past few decades, interdisciplinary courses have greatly increased (e.g., neuroscience, molecular biology, and environmental sciences; Lattuca & Voigt, 2022). An interdisciplinary course is organized around a topic, broadly defined as an issue, theme, or problem. Interdisciplinary courses offer an alternative to traditional knowledge production processes in that they seek to be integrative and holistic understandings of the social and natural worlds. An interdisciplinary pedagogy emphasizes an innovative approach to student-centered learning processes in developing their skills: evaluation, synthesis, integration, higher-order critical thinking, and problem-solving. Interdisciplinary courses allow students to develop the ability to tolerate multiple perspectives, to broaden horizons, and to increase willingness and capacity to question assumptions about the world and themselves.

We provided opportunities in this course for both interdisciplinary learning and working in teams to more effectively consolidate the students' understanding of sustainability. Interdisciplinary learning links across different subjects to enhance learning. Team learning involves working in groups to discuss critical issues, role plays, case studies, and sharing and understanding each other's views of the big picture. This interdisciplinary teaching methodology allowed students to learn complex issues related to food, energy, water, policy, and related social and economic issues in a novel and interactive manner. Figure 3 shows interactive learning in the classroom, and Figure 4 shows the students leading discussions.

Figures 3 and 4

Students Leading Discussions

Students were teamed in small groups. Students were encouraged to lead the discussion and also discuss their topics among themselves (interactive and peer-to-peer learning). The role of the instructors was to facilitate discussion within team members as well as among teams. Students were encouraged to discuss their topics among themselves (interactive and peer-to-peer learning).

Team-Building Games

In order to encourage learning in an enjoyable and retainable manner, we played several sustainability-themed games throughout the course. These included "Paying for Predictions," "Nuclear Jeopardy," and "Sustainability Jeopardy." These covered aspects of sustainability that were discussed throughout the course.

Grading and Projects

We determined the students' grades on the basis of their performance on a mid-term exam, a final exam, their in-class participation, and a group project. For the final project, the class was split into groups of four or six students, and each group was assigned a different project topic. Project titles included "Food Waste in Taiwan," "Plastic Waste in Taiwan and Policy," "Case Study: Water Pollution in Taiwan," "Circular Economy and Paper Industry in Taiwan," "Recycling Disposable Chopsticks in Taiwan," "Energy Policy in Taiwan," and "Climate Change in Taiwan." The goal of these group projects was to help students think critically about a given problem related to Taiwan, identify possible solutions, and then analyze the pros and cons of these solutions. At the end of the course the students were required to submit a project report and deliver a short team presentation. Figure 5 and Figure 6 represent pictures of students making presentations on "Food Security in Taiwan" and "Plastic Pollution in Taiwan."

Figures 5 and 6

Student Presentations on "Food Security in Taiwan" and "Plastic Pollution in Taiwan

Students were teamed in groups of four to six people. Teams were given a range of topics to work on team projects. These topics were cutting edge and key issues related to sustainability and environment, food, energy, water, climate. Teams were assigned time every day during the course period to interact with the teaching assistant and instructors to research their topics. Finally, students delivered a 15 minute team presentation to the whole class on the last day of the course: introduce topic, what did you research, what is the gap, what are possible technology solutions and policy recommendations to implement, and finally answer any questions from their peer students. Team presentations were graded on contents, and delivery. Teams also worked outside the class room to discuss their topic.

Data Analytics

Several analytical methods were applied in this study to examine the relationships between pretest and posttest for participants in their career interest in STEM and perceived knowledge. Zero-order correlations, analysis of covariance (ANCOVA), Fisher's r–to–z transformations and z-tests, and path analysis were utilized in data analytics. ANCOVA is a suitable analytical method since the data represented participants' STEM career interest and perceived knowledge in pre- and post-tests. Path analysis is used to reveal whether perceived knowledge at pre-test can be a significant predictor of career interest at post-test.

Results

Zero-order correlations are presented in Table 1.

Table 1

Zero-Order Correlations Among the Study Variables

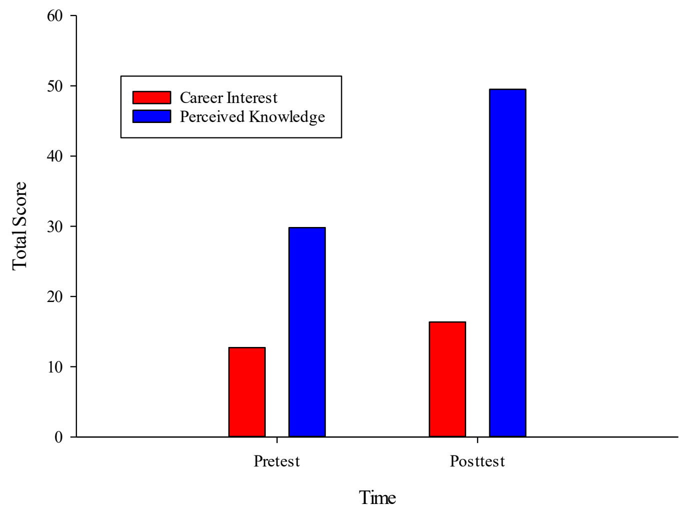
Variable	1	2	3	4
1. Pre-test career interest	—			
2. Post-test career interest	.50***	—		
3. Pre-test perceived knowledge	.19	.34*	—	
4. Post-test perceived knowledge	.09	.32*	.30*	—
5. *M*	12.72	16.36	29.82	49.50
6. *SD*	3.06	3.51	6.34	5.49

Note. * $p < .05$. *** $p < .001$

A bar graph of pre-test and post-test means for STEM career interest and perceived knowledge is also shown in Figure 7.

Figure 7

Pre-test and Post-test Means for STEM Career Interest and Perceived Knowledge



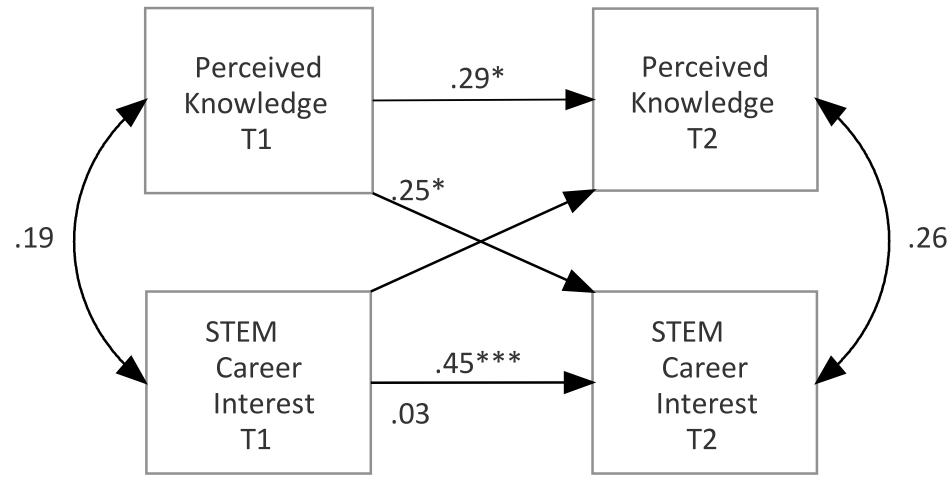
Pre-test career interest was positively and significantly correlated with post-test career interest ($r = .50$, $p < .001$), while pre-test perceived knowledge was also significantly associated with post-test perceived knowledge ($r = .30$, $p < .05$). Analysis of our four hypotheses is discussed below.

Our first hypothesis is that participants' perceptions of knowledge of environmental and sustainability increase significantly from pre-test to post-test. Our second hypothesis is that participants' individual interest in STEM careers will increase significantly from pre-test to post-test. To test the above two hypotheses, we conducted a repeated-measures analysis of covariance (ANCOVA) with time as the within-subject factor for both STEM career interest and perceived knowledge. We controlled for gender, program year, and the number of previous courses in which participants indicated they had learned about sustainability. Results revealed a significant multivariate effect of time (Pillai's $V = .50$, F (2, 39) = 19.48, $p < .001$, partial $\eta^2 = .50$). Despite this finding, follow-up univariate ANCOVAs indicated that only perceived knowledge increased significantly (F (1, 40) = 35.02, $p < .001$, partial $\eta^2 = .47$), from pre-test ($M = 29.82$, $SD = 6.34$) to post-test ($M = 49.50$, $SD = 6.34$). The effect of time for career interest approached significance, but it did not quite meet this threshold (F (1, 40) = 3.45, $p = .07$, partial $\eta^2 = .08$).

Our third hypothesis is that participants' knowledge of environmental sustainability at pre-test will be a significant positive predictor of STEM career interest at post-test. To test this hypothesis, we applied a cross-lagged panel regression model. To apply this model, we regressed the post-test interest and knowledge variables on their pre-test counterparts using path analysis. Results indicated that there were significant autoregressive effects, as pre-test interest was a significant positive predictor of post-test interest ($\beta = .45$, $p < .001$), and pre-test knowledge was a significant positive predictor of post-test knowledge ($\beta = .29$, $p = .04$). Importantly, perceived knowledge at pre-test was a significant positive predictor of career interest at post-test ($\beta = .25$, $p = .04$), thus supporting hypothesis three. See Figure 8 for information.

Figure 8

Cross-lagged Panel Model of STEM Career Interest Development



Our fourth hypothesis is that the positive correlation between post-test knowledge and post-test interest will be significantly stronger than the positive correlation between pre-test knowledge and post-test interest. To test this hypothesis, we converted the pre-test and post-test knowledge-interest correlations to z-scores using Fisher's r–to–z transformations and then performed a z-test on the difference between the two z-scores. Interestingly, the correlation between career interest and perceived knowledge was not significant at pre-test (r = .19, p = .21), but this relationship was significant at post-test (r = .32, p = .03). Nevertheless, results indicated that the difference between these correlations was not significant (z = -.33, p = .74); therefore, hypothesis four was not supported.

Discussion

The purpose of this study was to examine the efficacy of a novel college sustainability course in promoting relevant knowledge and interest in careers related to sustainability within a Taiwanese educational context for the purpose of raising awareness. Results from this study indicated that after participating in the course, students show improvement in their knowledge, interest, and self-efficacy over topics relating to sustainability.

The intervention was effective in increasing students' perceptions of knowledge, as their scores on the knowledge measure increased nearly 20 points from pre-test to post-test, supporting our first hypothesis regarding increased knowledge of environmental and sustainability issues. Our second hypothesis is not supported, as the increase in STEM career interest that was observed from pre-test to post-test—as predicted by the POI (person–object theory of interest) —was not statistically significant. The lack of a significant finding with respect to the second hypothesis regarding increased interest in STEM careers may be explained by the MDL (Alexander's model of domain learning). That is, the MDL states that individual interest should increase linearly as the learner's proficiency and knowledge structure becomes more coherent. Participants in our study may have developed subject matter proficiency over the course of the intervention, but one week affords little time to form a cohesive knowledge structure. However, information from multiple interdisciplinary fields relating to sustainability was introduced to the students during the course. Students were also required to participate in team discussions and at the end of the course, students delivered team presentations.

Hence, the amount of information students received and the coursework students had to do were considered intensive. If more time could be added to the course for other interactive activities such as field trips, the outcome of this short course would be even more effective, especially in terms of delivering a more cohesive knowledge to students.

The finding regarding hypothesis 3 is consistent with the MDL assertion that domain knowledge should provide a foundation for the development of interest over time. Exposure to a novel area of learning is a necessary precondition for students to develop the curiosity, attention, and enjoyment that is so critical to capturing interest in a task. Although we did not measure situational interest explicitly, it appears that greater understanding of the tasks that the participants were exposed to contributed to the formation of broader and more stable interest in STEM careers in general. This is consistent with previous findings which suggest that domain knowledge is positively associated with individual interest in STEM disciplines such as physics (Alexander et al., 1994; Tapola et al., 2013) and biology (Alexander et al., 1995; Durik & Matarazzo, 2009). The post-test correlation between knowledge and interest was not significantly stronger than the pre-test correlation between these variables, thus our fourth hypothesis was not supported. While the difference between the correlations was not statistically significant, our results are not inconsistent with the POI since there was a significant positive correlation between post-test career interest and post-test perceived knowledge. In our view, this represents preliminary evidence that the interaction of students with their learning environment fosters the development of interest. The lack of a significant finding may be due to insufficient statistical power given the relatively small sample size. Future research would do well to conduct a more robust test of this hypothesis with larger samples.

Our findings build upon the literature by demonstrating that knowledge within a unique environmental STEM domain is an important precursor to the development of interest in STEM careers. Why the participants expressed interest in these careers may be explained by their view of these careers being instrumental in helping them solve global challenges, such as climate change, food insecurity, and global health, or they may be interested in them for intrinsic reasons. In other words, they may see these careers as opportunities to achieve important personal or societal objectives, or they may simply be interested in STEM careers. Interestingly, our results showed no association between initial interest in STEM careers and greater acquisition of knowledge at a later point in time. The fact that initial knowledge was predictive of later interest, while the converse relationship (initial interest and later knowledge) was not significant, lends support to the MDL notion that proficiency in an academic domain is a necessary precondition for interest to develop.

Limitations and Future Research

As with any study, this one has limitations worth considering. As discussed in the results section, the sample was from a university in Taiwan. We recommend that future research examining the efficacy of similar courses be conducted with a larger sample of students from other cultures. We also did not directly measure participants' perceptions of situational interest following particular tasks. Thus, explicit connections between situational interest and STEM career interest could not be assessed. Future research evaluating the value of situational interest as a predictor of interest in STEM careers or even the more specific domain of environmental sustainability would be fruitful. Second, due to the lack of a control group, we cannot assume a causal relationship between the course and the observed effects. Although we statistically controlled for gender, academic classification, and prior knowledge (i.e., the number of previous courses in which students learned about sustainability), we could not make strong inferences of causality. Future research could build upon these results by including a control group or a comparison group that is exposed to a similar course or intervention. Furthermore, the efficacy of the intervention may have been limited by its selected short duration during fall break. Future research would do well to lengthen this or similar interventions to ensure that students' STEM career interests are sufficiently exposed to attitude-changing information. Furthermore, future research should also follow (e.g., a longitudinal study) the progress of students' career intentions throughout their years at university and may even include a post-graduation period to determine if their career choices are related to their initial interests in STEM subjects.

Future course offerings will aim to integrate other improvements into the curriculum and course structure. Students may be taken on tours of multiple industrial facilities to observe real world problems or document how they operate and affect sustainability. They will also be given the opportunity to work on industry-related projects by collaborating with local industrial partners. This will allow them to develop marketable skills and become competitive candidates when they decide to join the workforce. Another goal we have is to develop and expand the current course into a certificate or degree program in sustainability. Finally, to generalize our findings, it is important that future researchers reproduce this pilot study in other global settings and educational systems.

Conclusions

The four-day short sustainability development course introduces students to the environmental aspect through its interdependence with the nexus of energy, food, water, health, and economic considerations. The core content of the course described in this study covers the essential concepts of sustainability and introduces students to the environmental aspect through its interdependence with the nexus of energy, food, and water, health, and economic considerations. By creating an engaging class atmosphere with open-ended discussions and promoting experiential self-learning activities among the students, our findings suggest that this course was highly effective in enhancing the students' knowledge of key sustainability topics. Furthermore, the activities in the course such as group discussions, projects, and case studies helped to validate their understanding, broaden their perspectives, and allow them to appreciate the importance of the environment and sustainability.

This study's major contribution lies in the fact that it demonstrated the efficacy of a course that integrates environmental awareness in an engaging manner into STEM learning. It provided an opportunity for students to connect what they learn in the classroom to prominent socio-environmental concerns and issues of the real world, which they may gain a better understanding of, or even inspired to help solve, through such courses. An approach like this is needed, as some research has shown that sustainability in higher education sometimes fails to integrate environmental actions on campus with social and economic principles emphasized in community engagement projects (Zizka et al., 2021). In particular, Taiwanese university students are frequently limited in their understanding of sustainability issues, which they often see as only being related to individual actions or behavior, such as avoiding the use of straws, recycling plastic products, turning the lights off, and so on (Hsu & Pivec, 2021; Li & Liu, 2021). In particular, focusing on the nexus between energy, food, water and related economic and environment policy would help students gain a broader and more comprehensive understanding of modern sustainability issues.

This work was supported by the National Science and Technology Council in Taiwan under Grant #110-2511-H-005-001. The authors would like to thank encouragement by Tai Hui Lin and Rodney Matsuoka (both from National Cheng-Kung University, Tainan, Taiwan) and Maureen McCann (at Purdue University, West Lafayette, Indiana, USA).

Pankaj Sharma (sharma@purdue.edu) is a professor (courtesy) in the School of Engineering Technology at Purdue University (USA). He is also a project manager in the Laboratory of Renewable Resources Engineering, Department of Agricultural and Biological Engineering (Purdue University) and teaches online energy courses to MBA students. Earlier, Pankaj was associate director of

Discovery Park at Purdue University. He received his Ph.D. in physics from the Physical Research Laboratory, Gujarat University (India) and MBA from the School of Business, Purdue University (USA). He was a Fulbright New Century Research Scholar from 2009 to 2010. His research interests are in sustainability, geosciences, and STEM education.

Eric D. Deemer (edeemer@purdue.edu) is a professor in the Department of Educational Studies at Purdue University. He received his Ph.D. in counseling psychology from the University at Albany, State University of New York in 2008. His research focuses on academic and career development, with a particular emphasis on the motivational and psychological factors that foster the pursuit of careers in STEM.

Jane Lu Hsu (jlu@dragon.nchu.edu.tw) is a Distinguished Professor in the Department of Marketing at National Chung Hsing University in Taiwan. She received her Ph.D. in Agricultural Economics at Kansas State University. She is currently the Board Director of The Rural Economics Society of Taiwan, a prestigious academic association in agricultural and applied economics. She has been taking part in numerous national and international projects for agribusiness and environmental education and was awarded as a Fulbright New Century Scholar from 2009 to 2010. Her major research interests include data analytics, consumer marketing, and cross-cultural studies.

References

Aikens, K., McKenzie, M., & Vaughter, P. (2016). Environmental and sustainability education policy research: A systematic review of methodological and thematic trends. *Environmental Education Research*, *22*(3), 333–359.

Alexander, P. A. (2004). A model of domain learning: Reinterpreting expertise as a multidimensional, multistage process. In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 273–298). Routledge.

Alexander, P. A., Jetton, T. L, & Kulikowich, J.M. (1995). Interrelationship of knowledge, interest, and recall: Assessing a model of domain learning. *Journal of Educational Psychology*, *87*(4), 559–575.

Alexander, P.A., Kulikowich, J. M., & Schulze, S. K. (1994). How subject–matter knowledge affects recall and interest. *American Educational Research Journal*, *31*(2), 313–337.

American Nuclear Society. (2012). *Fukushima Daiichi: ANS Committee Report.* American Nuclear Society.

Argimon-Pallàs, J. M., Flores-Mateo, G., Jiménez-Villa, J., & Pujol-Ribera, E. (2011). Effectiveness of a short-course in improving knowledge and skills on evidence-based practice. *BMC Family Practice*, *12*(1), 1-7.

Borowy, I. (2021). The social dimension of sustainable development at the UN: From Brundtland to the SDGs. In C. Deeming (Ed.), *The struggle for social sustainability: Moral conflicts in global social policy* (pp. 89–108). Bristol University Press.

Breiting, S., & Mogensen, F. (1999). Action competence and environmental education. *Cambridge Journal of Education*, *29* (3), 349–353.

Brundtland, G. (1987). Report of the World commission on environment and development: Our common future (A/42/427). United Nations. Oxford University Press. Pages 383. ISBN 019282080X.

Bureau of Energy, Ministry of Economic Affairs (2022). *Energy Statistics Handbook 2021*. https://www.moeaboe.gov.tw/ECW_WEBPAGE/FlipBook/2021EnergyStaHandBook/index.html#p=.

Charatsari, C., & Lioutas, E.D. (2018). Environmental education in university schools: A study in a logistics faculty. *Applied Environmental Education & Communication*, *17*(2), 124–135.

Castillo, L., Gutierrez, W., & Gore. J. (2018). *Renewable energy saves water and creates jobs*. Scientific American. https://www.scientificamerican.com/article/renewable-energy-saves-water-and-creates-jobs/

Chen, S.-Y., & Liu, S.-Y. (2020). Developing students' action competence for a sustainable future: A review of educational research. *Sustainability*, *12*, 1374.

Choudhary, A., Saha, A. R., & Tiwary, N. K. (2019). The role of compulsory environmental education in higher learning: A study in the University of Delhi. *Applied Environmental Education & Communication*, *19*(4), 389–401.

Coyle, E., & Simons, R. (2014). *Understanding the global energy crisis*. Purdue University Press.

Deemer, E. D., & Sharma, P. (2019). Situational interest and scientific self–efficacy: Influence of an energy science career intervention. *The Career Development Quarterly*, *67*(2), 171–184.

Diamantoulakis, P., Kapinas, V., & Karagiannidis, G. (2015). Big data analytics for dynamic energy management in smart grids. *Big Data Research*, *2*(3), 94–101.

D'Odorico, P., Davis, K. F., Rosa, L., Carr, J. A., & Rulli, M. C. (2018). The global food–energy–water nexus. *Reviews of Geophysics*, *56*(3), 556–531.

Durik, A. M., & Matarazzo, K.L. (2009). Revved up or turned off? How domain knowledge changes the relationship between perceived task complexity and task interest. *Learning and Individual Differences*, *19*(1), 155–159.

Evans, N.S., Stevenson, R. B., Lasen, M., Ferreira, J.A. & Davis, J. (2017). Approaches to embedding sustainability in teacher education: A synthesis of the literature. *Teaching and Teacher Education*, *63*, 405–417.

Gunders, D. (2012). *Wasted: How America is losing up to 40 percent of its food from farm to fork landfill.* National Resources Defense Council.

Hazelton, P., Malone, M., & Gardner, A. (2009). A multicultural, multidisciplinary short course to introduce recently graduated engineers to the global nature of professional practice. *European Journal of Engineering Education*, *34*(3), 281-290.

Hidi, S. (2000). An interest researcher's perspective: The effects of extrinsic and intrinsic factors on motivation. In C. Sansone & J. M. Harackiewicz (Eds.). *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 309–339). Academic Press.

Hidi, S., & Harackiewicz, J. M. (2001). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research*, *70*(2), 151–179.

Hsu, J. L., & Pivec, M. (2021). Integration of sustainability awareness in entrepreneurship education. *Sustainability, 13*(9), 4934.

Jensen, B. B., & Schnack, K. (1997). The action competence approach in environmental education. *Environmental Education Research*, *3*(2), 163–178.

Judge, P. K., Buxton, J. A., Sheahan, T. C., Phetteplace, E. R., Kriebel, D. L., & Hamin Infield, E. M. (2020). Teaching across disciplines: A case study of a project-based short course to teach holistic coastal adaptation design. *Journal of Environmental Studies and Sciences*, *10*(3), 341-351.

Kaplan, J., Bond, T. J., Merson, M. H., Reddy, K. S, Rodriguez, K., Sewankamba, M. H., & Wasserheit, J. N. (2008). Towards a common definition of global health. *The Lancet*, *373*(9679), 1993–1995.

Kasimov, N., Malkhazova, S., & Romanova, E. (2002). The role of environmental education for sustainable development in Russian universities. *Planet*, *8*(1), 24–25.

Krapp. A. (2002). Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction*, *12*, 383–409.

Krapp, A., & Prenzel, M. (2001). Research on interest in science: Theories, methods, and findings. *International Journal of Science Education*, *33*(1), 27–50.

Lattuca, L.R & Voigt, L.J. (2022) *Interdisciplinary courses and majors in higher education*. Encyclopedia of Education. https://www.encyclopedia.com/education/encyclopedias-almanacs-transcripts-and-maps/interdisciplinary-courses-and-majors-higher-education.

Lawless, K. A., & Kulikowich, J. M. (2006). Domain knowledge and individual interest: The effects of academic level and specialization in statistics and psychology. *Contemporary Educational Psychology 31*(1), 30–43.

Lehman, P. K., & Geller, E. S. (2004). Behavior analysis and environmental protection: Accomplishments and potential for more. *Behavior and Social Issues*, *13*(1), 13–32.

Li, Y. Y., & Liu, S. C. (2021). Examining Taiwanese students' views on climate change and the teaching of climate change in the context of higher education. *Research in Science & Technological Education*, 1–14.

Luo, X., Wang, J., Dooner, M., & Clarke, J. (2015). Overview of current development in electrical energy storage technologies and the application potential in power system operation. *Applied Energy*, *137*, 511–536.

McKinsey Center for Business and Environment. (2016). *The circular economy: Moving from theory to practice*. McKinsey Center for Business and Environment.

Miller, H., Miller Jr, B. R., & Spoelstra, J. (2021). A sustainability internship program: Strategies for creating student stewards for sustainability. *International Journal of Sustainability in Higher Education* (ahead of publication). https://doi.org/10.1108/IJSHE-08-2020-0314

Murphy, P. K., & Alexander, P. A. (2002). What counts? The predictive powers of subject-matter knowledge, strategic processing, and interest in domain-specific performance. *The Journal of Experimental Education, 70*(3), 197-214.

Misra, A. (2014). Climate change and challenges of water and food security. *International Journal of Sustainable Built Environment*, *3*(1), 153–165.

National Academy of Engineering & National Research Council (2012*). Macondo well eeepwater Horizon blowout: Lessons for improving offshore drilling safety*. National Academy of Engineering.

National Research Council. (2001). *Grand challenges in environmental sciences*. The National Academies Press.

Olsson, D., Gericke, N., Sass, W., & Boeve-de Pauw, J. (2020). Self-perceived action competence for sustainability: The theoretical grounding and empirical validation of a new research instrument. *Environmental Education Research*, *26*(5), 742-760.

Owusu, G.M.W., Kwakye, T.O., Welbeck, E. E., & Ofori, C.G. (2017). Environmental literacy of business students in Ghana. *International Journal of Sustainability in Higher Education*, *18*(3), 1467–6370.

Qablan, A., Southerland, S. A., & Saka, Y. (2011). "My job isn't to tell them what to think": The fear of indoctrination and how it shapes education for sustainable development. *Electronic Journal for Research in Science & Mathematics Education*, *15*(2), 1–28.

Redclift, M. (2006). Sustainable development (1987–2005): An oxymoron comes of age. *Horizontes Antropologicos*, *13*(4), 65–84.

Rockström, J., Steffen, W., Noone., K., Persson, A., Chapin III, F., Lambin, E., & Karlberg, L. (2009). A safe operating space for humanity. *Nature*, *461*(7263), 472–475.

Sass, W., Boeve-de Pauw, J., De Maeyer, S., & Van Petegem, P. (2021). Development and validation of an instrument for measuring action competence in sustainable development within early

adolescents: The action competence in sustainable development questionnaire (ACiSD-Q). *Environmental Education Research, 27*(9), 1284-1304.
Scott, W. (2002). Education and sustainable development: Challenges, responsibilities, and frames of mind. *The Trumpeter*, *18*(1), 1-10.
Seidelin, L., Albrechtsen, T. R. S., Schöps, K., Holmer, M., & Wahlberg, M. (2021). Does teaching about artificial reefs trigger students' situational interest in marine biology? *Journal of Biological Education, 55*(3), 264-275.
Şterbuleac, D., & Toma, O. (2019). The overlooked role of academic environmental organizations in addressing environmental education issue. *Applied Environmental Education & Communication*, *18*(1), 91–94.
Şterbuleac, D., & Toma, O. (2020). Environmental education through a university eco–consortium. *Applied Environmental Education & Communication*, *19*(1), 62–73.
Taiwan Ministry of Agriculture. (2023). *Agricultural statistics yearbook*. Executive Yuan.
Tapola, A., Veermans, M., & Niemivirta M. (2013). Predictors and outcomes of situational interest during a science learning task. *Instructional Science*, *41*(6), 1047–1064.
The World Bank. (2018, September 28). *Nam Theun2 Project Overview*. The World Bank. http://www.worldbank.org/en/country/lao/brief/nam–theun–2–project–overview–and–update
Tilbury, D. (2011). Education for sustainable development: An expert review of processes and learning (ED–2010/WS/46). UNESCO.
Tröbst, S., Kleickmann, T., Lange–Schubert, K., Rothkopf, A., & Möller, K. (2016). Instruction and students' declining interest in science: An analysis of German fourth- and sixth-grade classrooms. *American Educational Research Journa*l, *53*(1), 162–193.
Ufer, S., Rach, S., & Kosiol, T. (2017). Interest in mathematics = Interest in mathematics? What general measures of interest reflect when the object of interest changes. *ZDM Mathematics Education*, *49*(3), 397–409.
United Nations Report. (2015). *The millennium development goals report 2015*. United Nations. http://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20Summary%20web_english.pdf
United Nations Report. (2016). *The sustainable development goals report 2016*. United Nations. http://www.globalamalen.se/wp–content/uploads/2016/05/The–Sustainable–Development–Goals–Report–20161.pdf
U.S. Energy Information Administration. (2018). *Use of energy in the United States explained*. U.S. Energy Information Administration. https://www.eia.gov/energyexplained/index.php?page=us_energy_use
Vennix, J., den Brok, P., & Taconis, R. (2018). Do outreach activities in secondary STEM education motivate students and improve their attitudes towards STEM? *International Journal of Science Education, 40*(11), 1263-1283.
Zizka, L., McGunagle, D. M., & Clark, P. J. (2021). Sustainability in science, technology, engineering and mathematics (STEM) programs: Authentic engagement through a community-based approach. *Journal of Cleaner Production*, *279*(2021), 123715.

Appendix I

Course Schedule: Sustainability for Future Generations - Environmental, Economical, Managerial, and Health Perspectives

September 3rd (Monday)	
12:00-1:00 PM	Luncheon with Faculty and Students
1:00-1:30 PM	Free Time
1:30-2:50 PM	Introduction
2:50-3:00 PM	Break
3:00-4:20 PM	Sustainability Perspective
4:20-5:00 PM	Project Discussion with Instructor
5:00-7:00 PM	TA's Session: Work on Group Project
September 4th (Tuesday)	
10:30 AM -12:00 PM	Energy Perspective: Fossil and Renewable Resources
12:00-1:00 PM	Lunch
1:00-1:30 PM	Free Time
1:30-2:50 PM	Sustainable Agriculture Perspective: Food and Water
2:50-3:00 PM	Break
3:00-4:20 PM	Technology Perspective: Environment and Climate Change
4:20-5:00 PM	Project Discussion with Instructor
5:00-7:00 PM	TA's Session: Work on Group Project
September 5th (Wednesday)	
10:30 AM -12:00 PM	Managerial Perspective: Policies, Politics, and Regulations
12:00-1:00 PM	Lunch
1:00-1:30 PM	Free Time
1:30-2:50 PM	Business and Economics Perspective: Technology Diffusion and Circular Economy
2:50-3:00 PM	Break
3:00-3:55 PM	Global Health Perspective: Life-Style Changes and Diseases
355-4:05 PM	Break
4:05-5:00 PM	NEXUS Perspective: Interdependencies and Interconnections
5:00-7:00 PM	TA's Session: Work on Group Project
September 6th (Thursday)	
Decide your own time	Work in Your Group for Presentations
3:30-5:00 PM	Group Presentations
5:00-5:30 PM	Prizes, Certificate Presentations, and Group Photo
5:30-7:00 PM	Dinner with Faculty, Dean, TA, and Others
7:00 PM	Good Bye

Appendix II

STUDENT PRE-PARTICIPATION SURVEY

1. Name ____
2. Nationality ______________________________
3. Gender _______________
4. In what year were you born? _______________
5. Year in School: ______________________________
6. Which of the following courses did you take in the past 2-3 years (Select all that apply)?

Biology	Chemistry	Mathematics
Earth Science	Economics	Physics
Engineering	Technology	Other (please specify) __________

7. In which of the following courses have you learned about energy, water, food, environment and health related topics?

Biology	Chemistry	Mathematics
Earth Science	Economics	Physics
Engineering	Technology	Other (please specify) __________

8. How would you rate your interest in the following careers?

Please circle the most appropriate answer for each statement	Never considered	Very weak	Moderate	Very strong
Math-related career (e.g., accountant, economist,)	1	2	3	4
Science-related career (e.g., biologist, chemist, geologist).	1	2	3	4
Engineering related career.	1	2	3	4
Technology or computer related career.	1	2	3	4
Environment related career.	1	2	3	4

9. What do you know about sustainability?

Not Much				A Great Deal
1	2	3	4	5

Comments:

10. What do you know about energy

Not Much				A Great Deal
1	2	3	4	5

Comments:

11. What do you know about Climate Change?

STUDENT PRE-PARTICIPATION SURVEY

Not Much				A Great Deal
1	2	3	4	5

Comments:

12. What do you know about food?

Not Much				A Great Deal
1	2	3	4	5

Comments:

13. What do you know about water?

Not Much				A Great Deal
1	2	3	4	5

Comments:

14. What do you know about environment?

Not Much				A Great Deal
1	2	3	4	5

Comments:

15. What do you know about the economics of energy?

Not Much				A Great Deal
1	2	3	4	5

Comments:

16. What do you know about policy, politics, regulation, and strategy?

Not Much				A Great Deal
1	2	3	4	5

Comments:

17. What do you know about the nexus of energy, food, water, climate, and environment?

Not Much				A Great Deal
1	2	3	4	5

Comments:

18. What do you know about the circular economy?

Not Much A Great Deal

STUDENT PRE-PARTICIPATION SURVEY

1	2	3	4	5

Comments:

19. What do you know about the global health?

Not Much				A Great Deal
1	2	3	4	5

Comments:

20. What do you hope to learn (or gain) from this course?

Appendix III

STUDENT POST-PARTICIPATION SURVEY

1. Name ____

2. How would you rate your interest in the following careers?

Please circle the most appropriate answer for each statement	Never considered	Very weak	Moderate	Very strong
Math-related career (e.g., accountant, economist,)	1	2	3	4
Science-related career (e.g., biologist, chemist, geologist).	1	2	3	4
Engineering related career.	1	2	3	4
Technology or computer related career.	1	2	3	4
Environment related career.	1	2	3	4

3. What do you know about sustainability?
   Not Much A Great Deal
   1 2 3 4 5

   Comments:

4. What do you know about energy?
   Not Much A Great Deal
   1 2 3 4 5

   Comments:

5. What do you know about Climate Change?
   Not Much A Great Deal
   1 2 3 4 5

   Comments:

6. What do you know about food?
   Not Much A Great Deal
   1 2 3 4 5

   Comments:

7. What do you know about water?
   Not Much A Great Deal
   1 2 3 4 5

   Comments:

8. What do you know about Environment?
   Not Much A Great Deal
   1 2 3 4 5

STUDENT POST-PARTICIPATION SURVEY

Comments:

9. What do you know about the economics of energy?

Not Much				A Great Deal
1	2	3	4	5

Comments:

10. What do you know about policy, politics, regulation, and strategy?

Not Much				A Great Deal
1	2	3	4	5

Comments:

11. What do you know about the nexus of energy, food, water, climate and environment?

Not Much				A Great Deal
1	2	3	4	5

Comments:

12. What do you know about the circular economy?

Not Much				A Great Deal
1	2	3	4	5

Comments:

13. What do you know about the global health?

Not Much				A Great Deal
1	2	3	4	5

14. What do you know about the natural and man-made disasters?

Not Much				A Great Deal
1	2	3	4	5

Comments:

15. What did you think of the course? What were your major take-aways?

16. What are some things you liked about the course?

STUDENT POST-PARTICIPATION SURVEY

17. What are some things you would suggest changing about the course?

18. Describe – how course was inspiring to think like a leader?

ELECTRONIC JOURNAL FOR RESEARCH
IN SCIENCE & MATHEMATICS EDUCATION
VOL. 27, NO. 2, 122-157

Predicting Preservice Teachers' Performance on the Science Core of the EC-6 TExES General Certification Examination

Ruthanne Thompson
University of North Texas

Naudin Alexis
University of North Texas

Pamela Esprivalo Harrell
University of North Texas

ABSTRACT

Predicting preservice teachers' performance on their certification examination may meaningfully help Educators Preparation Programs (EPPs) to adapt and integrate learning frameworks that can improve their passing rates. This study used multiple linear regression (MLR) and binomial logistic regression (BLR) to explore potential variables that may impact the preparedness of 170 pre-service teachers to pass the science core of the official EC-6 Texas Examinations for Educator Standards (TExES) certification examination. The study was conducted by issuing a practice EC-6 TExES certification examination in a pretest and post-test manner during the semester that the participating cohort were enrolled in BIOL 1082, a mandatory science course EC-6 preservice teachers need to take prior to the official state EC-6 TExES certification exam. Additionally, the cohort took an online Qualtrics™ survey that collected ex post facto and other demographics data. The independent variables explored in this study included: final grade in BIOL 1082, classification, transfer status, prior college science coursework, enrollment status, family's college history, and current GPA. The dependent variable used was the post-test score on the practice EC-6 exam. The independent variable, grade in BIOL 1082 was revealed to be the single best predictor of preservice teachers' performance on the science practice examination across both the MLR and BLR models. The BLR models had a higher prediction accuracy of preservice teachers who would most likely fail the practice test than those who may pass at a prediction rate at approximately 79% accuracy. Based on the 67 out of 170 preservice teachers who passed the post-test, the accuracy of predicting failures may be a useful tool that EPPs can use in identifying students who may be at risk of failing and thus implement necessary interventions and other educational strategies.

Keywords: preservice teachers, multiple linear regression, binomial logistic regression, predictor factors, certification

Introduction

Science and mathematics have long been considered the toughest subjects for students to master at the primary level through the tertiary level of education (Murphy & Smith, 2012; Pino-Pasternak &

Volet, 2018; van Aalderen-Smeets et al., 2017). The belief in preconceived poor performance in science have at times caused even the most brilliant of students to pivot to courses and degrees that were considered less challenging (Pino-Pasternak & Volet, 2018). The carried belief of mediocre performance in science at times has come from family members and for some students their low self-efficacy in science has come from their own teachers (Pino-Pasternak & Volet, 2018). Further, according to Murphy et al., (2007) many preservice teachers have themselves predicted that they will not perform well on the science portion of their certification examination. Additionally, some elementary classroom teachers, even after passing their certification examination, have expressed that they spend the least amount of time on the subject matter of science in the classroom (Binns et al., 2020). This cyclical apprehension of preservice teachers' ability to learn and teach science can in turn lead to their future students getting low exposure to science content, which can then be reflected as an ongoing lapse in students' commitment to persevere in learning science. Later, if these same students pursue teacher certification, they may unintentionally relay to their students that success in science is unachievable (Binns et al., 2020; Pino-Pasternak & Volet, 2018). Studies have also shown that a connection exists between the passing of the certification exam and a preservice teacher's effectiveness in the classroom (Boyd et al., 2007; Donovan et al., 2014; Senler, 2016).

In the state of Texas, to obtain an elementary teacher certification one must pass the Texas Examination for Educator Standards (TExES) Early Childhood to 6th grade (EC-6) examination (EC-6 TExES). Prior to 2015, students were allowed an unlimited number of attempts to pass the exam. After 2015, with the passing of the *EESA Act*, examinees are only allowed five attempts to pass the exam with a score of at least 80%. In tandem with this shift in the number of attempts allowed, the Educator Preparation Programs (EPPs) must also maintain an overall pass rate of 85% each year to meet accreditation requirements (Warren, 2017). Creating a model to predict preservice teachers' performance on their official certification examination may not only help boost a preservice teacher's motivation, self-efficacy, and confidence (Ebrahim, 2012), but may also serve as a beneficial tool to the EPPs (Hutson, 2017).

The TExES™ EC-6 certification examination consists of two subdivisions: content area for the specific grade level, and pedagogy and professional responsibilities (Feuer et al., 2013; Gard, 2011; Sasson, 2014). There are five subject domains within the content portion of the exam: Domain I: English Language Arts Reading, & the Science of Teaching Reading (28%); Domain II: Mathematics (18%); Domain III: Social Studies (16%); Domain IV: Sciences (19%); and Domain V: Fine Arts, Health, and Physical Education (19%) (TExES™ program preparation manual, 2019). This study focused on subject Domain IV: Sciences, which includes 18 competencies over which 52 multiple choice questions are asked on the official EC-6 TExES exam. The courses offered at the EPP prepare preservice teachers for the science portion of the EC-6 certification examination and include BIOL 1082 (Biology for Elementary Educators), BIOL 1132 (Environmental Science for non-science majors), GEOG 1710 (Earth Science for non-science majors), PHYS 1210 (Conceptual Physics for non-science majors), and EDEE 4330 (Science Methods for Elementary Educators). The Biology for Educators course (BIOL 1082) was developed exclusively for preservice Elementary Education majors and covers approximately 8 of the 18 competencies of Domain IV, which makes up about 44% of the science concepts that need to be mastered for the certification examination. The other three science courses, Environmental Science, Earth Science, and Conceptual Physics are also taught within the College of Science; however, these courses are open to all non-science majors and are not taught in context of becoming an elementary teacher. In their senior year, elementary education majors, take EDEE 4330, Science Methods, which is taught within the College of Education. Since Biology for Elementary Educators has content that covers material from the other three science courses and, as a single course, covers the largest amount of the material on the official examination (44%), it was selected as the focus for this study.

Predicting preservice teacher passing rates on the state certification examination and identifying potential variables that may influence their performance is beneficial to EPPs, preservice teachers themselves, and other agencies (Hutson, 2017). Various studies have explored the impact of academic and environmental factors that may influence students' performance on the teacher certification examination. Most studies have included academic factors such as GPA, familial influences, age, workload, full time or part-time status, and other environmental factors (DeFreitas & Rinn, 2013; Graunke & Woosley, 2005; Kim & Corcoran, 2018; Peters & Draughon, 2017; Swecker et al., 2013). Though studies have explored factors that may impact or predict performance on preservice teachers state certification examination, most of these studies looked at how preservice teachers performed on the overall content exam where all the subject areas were combined, or their performance was examined based on performance on the Pedagogy and Professional Responsibilities (PPR) with factors such as GPA, first-generation, and course grades predicting performance on the overall certification examination (Frizzell, 2014; Gard, 2011; Kazempour & Sadler, 2015; Kim & Corcoran, 2018; Warren, 2017).

Of the few studies that have explored preservice teachers' performance on individual subject matter exams, academic factors showed some correlation to predicted performance of preservice teachers on the certification examination. In previous studies, exploring predictors of performance on teacher certification examination, Gard (2011) investigated factors predicting failure on TExES 8-12 History and found that transfer status and GPA were closely correlated to success on the examination. Thobega and Miller (2008) explored factors predicting preservice teachers' performance in agriculture certification and revealed that ACT scores and gender were highly correlated to success on the PRAXIS II content examination. On the other hand, Sandholtz and Shea's (2015) exploration of preservice teachers' predicted performance in California state licensing revealed that grades in prerequisite courses as well as supervisors anticipated performance of preservice teachers were not accurate predictors of performance on certification examinations.

Further, the National Council on Teacher Quality (NCTQ) stated that almost half of EPPs do not analyze preservice teacher content knowledge before and after science courses, nor do they associate such coursework with predicted/actual performance on certification examinations (NCTQ, 2014). In fact, most of the previous studies exploring teacher certification examinations were in subject areas of mathematics or history or addressed the overall general content examination. In the single study that attempted to dissect the subject matter within the content area of EC-6 generalists (Bains, 2011), the trajectory of the study was to investigate the pass rate of each subject area. Thus far, no study has been identified that has exclusively explored the predictive modeling of preservice teachers' performance on the science portion of the content area of certification examinations. Thus, there remains a gap in the literature on investigating factors that may affect preservice teachers' performance on the individual subject domains within the EC-6 TExES content examination. With science being a subject matter that seems daunting to both preservice teachers and their future students, predicting performance of preservice teachers on science subject area (Domain IV) of the EC-6 TExES certification exam was targeted as the subject matter to be explored (Warren, 2017). Based on factors from previous studies and a class survey, variables showing correlation to students' performance were selected as independent variables and analyzed using multiple linear regression (MLR) and binomial logistic regression (BLR). Variables such as grades, transfer status, previous courses taken, credits taken, first-generation and relatives in education were identified using a survey. The outcome variable was the score on the science practice exam.

EPPs can lose their accreditation status and preservice teachers can accrue additional debt and lowered self-confidence if failure on the official EC-6 TExES certification exam occurs (Hutson, 2017). Additionally, the rigor for each certification examination has increased along with an increase in student population in all classes in every state (Darling-Hammond, 2019). The findings of this study may help guide EPPs on how and where to best utilize their resources to support preservice teachers prior to their taking of the official certification examination and may also help guide preservice teachers on which science competencies they need to most focus their remediation efforts. Additionally, predictive factors identified as having a potential impact on performance may be useful in the early part of the recruitment of preservice teachers into EPP programs. (Linn & Jacobs, 2015).

Literature Review

Learning and teaching science is still regarded as one of the most difficult tasks for preservice teachers to grasp (Hutson, 2017). In studies assessing preservice teachers' readiness to teach science in a classroom, approximately 70% of first-time classroom teachers expressed nervousness and unpreparedness to teach science in comparison to teaching the other subjects (Binns et al., 2020; Pino-Pasternak & Volet, 2018). Pre ESSA Act (2015), pre-service teachers could do extremely well in one subject area and that would compensate for doing poorly in another, so long as the overall score was passing. Post *ESSA* Act, minimal proficiency must be achieved in all subjects as each is scored separately and each must reach the 80% threshold. For this reason, having a solid knowledge on each of the 18 competencies for Domain IV is essential (Sawyers & Myers, 2018). Predictive factors that can influence preservice teachers' performance on their official certification examination can be a powerful tool used by EPPs to focus their attention and resources on the areas and factors to help in both the retention of enrollees within their program as well as improving the pass rate of their students (Warren, 2017). Preservice teachers that score at or below the required 80% on their practice examination can receive scaffolding and other types of interventions to help improve their scores on the official examination. Additionally, a predictive model can highlight the independent variables, both academic and environmental, that may predict actual performance.

Regression models have long been used to build predictive models, with multiple linear regression being the most used (Sullivan et al., 1996). In this study, both multiple linear regression, and binomial logistic regression models were used to predict performance. While a binomial logistic model can be employed to predict broad dichotomous outcomes such as *pass* or *fail*, multiple linear regression offers a closer estimation of the actual score. For instance, logistic regression can place individuals into a *pass* category, however the score could be 79.9%, which would be on the cusp of failing. Thus, multiple linear regression provides the detailed score, allowing the EPP and learner to know that they can still be in danger of failing.

In 2002, the *NCLB Act*, shuffled the deck to reform America's education system. One of the changes of the *NCLB Act* was increased rigor for teacher certification examinations (Darling-Hammond, 2019). Plans and policies were written at the federal level. For example, schools were evaluated based on their Annual Yearly Progress (AYP) and sanctions were issued if AYP goals were not met for three consecutive years (Miller & Hudson, 2007; Lesley, 2016). From 2012 to 2015, the Obama administration revised parts of the *NCLB Act* and in 2015 proposed the *Every Student Succeeds Act* (*ESSA*). *ESSA* (2015) discarded the AYP requirements and returned the accountability back to the state level where requirements became evidenced-based measures to improve standards for all students (Darling-Hammond et al. 2016). For Texas, this translated to the TExES and under the TExES™ EC-6 test Domain IV, each of the competencies outline the standards for which the preservice teacher must demonstrate proficiency to successfully teach EC-6 science (TEA, 2018). Studies have shown that preservice teachers usually score poorly in science Domain IV (Miller & Hudson, 2007; Kazempour & Sadler, 2015).

One of the critical changes that came along with the *ESSA Act* of 2015 was limiting the number of attempts preservice teachers can take certification examination. A second change targeted how the examination was scored and required each of the five subject domains to be passed with a minimum score of 80% rather than using the former composite score which allowed low performance in some subject domains to be offset by higher performance in other domains (Darling-Hammond et al., 2016; Hutson, 2017). The reduction in number of attempts, along with increased pressures to improve performance of students in both mathematics and science placed additional stress on preservice teachers to master all subject domains, and on EPPs to improve performance on certification examinations (Darling-Hammond et al., 2016; Sutcher et al., 2016).

In addition to changes to the structure and scoring of the certification exams, there continues to be a nationwide shortage of teachers (Sutcher et al., 2016). This shortage has been due to an increase in the population of students attending school; an increase in the number of retirees; a decrease in the number of preservice teachers pursuing certification by approximately 33%; a decrease in the passing rate of teachers on their certification exam; and an increase in the standards across all platforms that measures teachers' accountability. The EPPs that prepare preservice teachers can also be sanctioned or receive disciplinary actions if their program fails to produce high quality teachers (Warren, 2017).

Research on determining factors that can affect and predict student performance on examinations continues to be relevant and various statistical methods and academic and environmental factors have been investigated to determine impact (Kazempour & Sadler, 2015; Kim & Corcoran, 2018). For example, Lykourentzou et al., (2009), using a multiple linear regression model, found that the best prediction of performance in an online course was the staging of practice multiple choice examinations early in the course. D'Amico and Dika (2013) examined precollege and other academic factors during early college years that may influence students' success in Science, Technology, Engineering, and Mathematics (STEM) fields and found GPA and success in mathematics were major indicators of success on the examination. Kim and Corcoran (2018) investigated factors that impact preservice teacher academic achievement and showed GPA, along with performance in their EPP, were impactful on a student's test performance. Studies investigating the impact of age and college classification on examination performance revealed that older students, as well as those who are further along in their college classification, showed increased cognitive skills as well as higher GPA (Kim & Corcoran, 2018). Also evident, was that factors investigated, in terms of college classification, focused on the performance of freshmen in college (Graunke & Woosley, 2005). Additionally, similar to freshmen, their study revealed that sophomores were at significant risk of dropping out of college. Furthermore, first-generation students tend to earn lower grades and have a lower completion rate of college compared to the other college counterparts as they are often employed whilst attending college (DeFreitas & Rinn, 2013; Kim & Corcoran, 2018; Martinez et al., 2009; Swecker et al., 2013). In addition, transfer students, compared to non-transfer students, had lower GPAs however, the GPA advantage held by non-transfer students disappeared about a semester after the student transferred (Douglass, 2012; Krieg, 2010). Finally, over the last five decades, increasing numbers of college students engage in part-time work, with approximately 40% of all full-time students having a part-time job (Peters & Draughon, 2017). Since it is anticipated that by 2025 close to 45% of all college students will be attending part-time, the influences on part-time students and what determines their success in college needs to be investigated (NCES, 2017; Peters & Draughon, 2017). With all of these independent variables shown to have some impact on student performance, this study used some of these variables to evaluate their individual, as well as their combined, impact on performance on the science portion of the EC-6 TExES certification examination.

Predicting preservice teachers' performance on their certification examination with the use of predictive models can allow EPPs to identify students who may potentially fail the official state certification examination so they may implement strategies to improve success on a subsequent EC-6 certification exam through remediation, retrieval practice, improved test taking strategies, and other interventions (Masters, 2018). This study used BIOL 1082, a mandatory preservice biology course that covers the biology competencies of the EC-6 TExES examination. Students who volunteered to participate in the study took the practice EC-6 certification examination at the beginning of the semester (pretest) and at the end of the semester (post-test). Additionally, an online Qualtrics™ survey was used to collect self-reported demographics and other pertinent information about the participants. In support of the goal of this study, which was to predict preservice teachers score on the science portion of the EC-6 TExES practice certification exam, the grade in BIOL 1082 as well as 13 additional independent variables from the Qualtrics survey were utilized. The post-test score on the practice exam was used as the dependent variable. Taken together, these variables were used to create multiple linear regression and binomial logistic regression predictive models. These predictive models can arm EPPs with tools to help them identify preservice teachers in need of scaffolding and remediation for content knowledge and will help the preservice teachers maintain and or achieve acceptable levels of academic success. Given this goal, this study set out to answer the following research question:

> Do any of the following explanatory variables: final grade in BIOL 1082 (X1), classification (X2), transfer status (X3), college biology (X4), college chemistry (X5), college physics (X6), college environmental science (X7), college earth science (X8), part-time (X9), credits taking (X10), first-generation college student (X11), relatives with degree in education (X12), and current GPA (X13), individually or in combination predict preservice teachers' performance on the science portion of the practice EC-6 TExES™ examination?

Methods

This study focused on preservice teachers enrolled in the Biology for Educators class (BIOL 1082), a mandatory course for preservice elementary educators, taught at a university in the southwest area of the United States. The participants took a Qualtrics™ survey to identify demographics as well as independent variables associated with performance. The dependent variable was the post-test score on the practice EC-6 TExES certification examination. MLR and BLR were used to develop predictive models associated with student success on the certification examination.

Participants

This study collected data from 170 preservice teachers pursuing a Bachelor of Science in Education with EC-6 teacher certification in the College of Education. These teachers were enrolled in Biology for Elementary Educators (BIOL 1082), within the College of Science, over the span of three consecutive semesters. The course BIOL 1082 can be taken at any time during the program therefore, all four classifications of students: freshman, sophomore, junior and seniors were participants in this study. The only science course for EC-6 TExES certification that had a set time frame for when it can be taken was EDEE, which was only available during the participants senior year. The preservice teachers age ranged from 18 to 35 years old. Most participants (62.4%) were age 18-20. 28.8% were between 21-23 years of age, 6.47% were 24-26, 1.17% were 27-30, and 1.17% were age 31-35. The classification of the preservice teachers was a mixture of all four classification levels with 21.8% freshman, 32.9% sophomore, 37.6% juniors and 7.64% seniors. Approximately 90% of the participants were full-time students. Half of the preservice teachers were transfer students. The participants of this study were predominantly female (97%) which is typically the case seen with the decline of male preservice teachers in early childhood education (Stroud et al., 2006), and is consistent with the general gender distribution for the EC-6 major at this University. The survey was used to identify science coursework completed during high school. The survey showed that 97% took biology, 93.5% took chemistry, 88.2% took physics, 20% took environmental science, and 8.2% took earth science. The participants were also asked if they took biology, chemistry, physics, environmental science, and earth science at college to which they responded: 28.8% took biology, 8.2% took chemistry, 30.6% took physics, 51.2% took environmental science, and 38.8% took earth science. Approximately 65.3% of the participants had a relative who had obtained a degree in education. Only 113 out of 170 (66.5%) of preservice teachers reported their overall GPA of which 5.3% had GPAs of 2.0-2.5, 14.2% had GPAs of 2.51-3.0, 37.3% had GPAs of 3.01-3.5, and 43.4% had GPAs of 3.15-4.0. Approximately 58.1% of preservice teachers reported being first-generation college students. Of the participants, 168 out of 170 (98.8%) preservice teachers completed BIOL 1082 of which 159 out of 170 (93.5%) passed BIOL 1082 with a Grade "C" or better.

Data Collection

Preceding the start of the study, authorization to conduct this research was asked for from the Institutional Review Board (IRB). Preservice teachers who volunteered to participate in the study were required to sign the IRB (Code No. 17-206) consent form before participating. Quantitative and qualitative data was collected from 170 preservice teachers. An online survey (Biology for Educators) was used to collect ex post facto qualitative and quantitative information from preservice teachers which included demographics such as age, classification, previous science courses taken, first generation, part-time or fulltime student status, number of credits already taken, transfer vs non-transfer, and relatives with degrees in education.

A practice test on the science core of the EC-6 TExES™ certification examination was prepared and issued by the TExES™ Advising Office (TAO). The practice EC-6 TExES™ exam was given during the second week of the semester during a class period of BIOL 1082 and labeled pretest. The practice exam is a secured examination developed for the state certification agency and consisted of 45 multiple choice questions asked over the 18 competencies that preservice teachers need to be proficient in over Domain IV of the EC-TExES certification exam. Preservice teachers documented their answers on scantrons, which were evaluated and assessed by the TAO, and their coded responses were returned to the researcher. The coded responses included the 18 competencies and the total number of questions asked over every competency and the number of questions each preservice teacher got correct over that competency and can be seen in Table X. The practice test was given again to the students as a post-test the second to last week of that same semester.

The BIOL 1082 course specifically addressed competencies: C1, C2, C4, C10, C11, C12, C13, and C14 which are the competencies for biology for educators. The preservice teacher degree program does not require a chemistry course and competencies C8 cover topics associated with a chemistry course although they may be taught during a physics course as well (e.g., waves, periodic table). The PHYS 1210 course addressed competencies: C7, C8, and C9 in the subject area of physics. The BIOL 1132 course addressed competencies: C14, and C17 in the subject area of environmental science. The GEOG 1710 course addressed competencies: C15, C16, C17, and C18 in the subject area earth science. The EDEE 4330 course addressed competencies: C3, C5 and C6 and covers the scope and sequence of science education from early childhood to 6th grade. For the practice EC-6 TExES certification exam the number of questions asked over each of the 18 competencies are summarized in Table 1.

Table 1

Number of Questions asked over each Competency of Domain IV EC-6 TExES Certification Examination

Competencies (C) 1-18	Number of questions out of 45 questions asked
C1	3
C2	3
C3	2
C4	2
C5	3
C6	2
C7	3
C8	3
C9	2
C10	2
C11	3
C12	1
C13	2
C14	3
C15	3
C16	3
C17	2
C18	3

Description of Independent Variables and Prediction Models

The regression models used a total of 13 factors, or independent variables, to develop four multiple regression models and four binomial logistic models. For ease of representation, the independent variable will be referred to as follows: grade in BIOL 1082 (X1), classification (X2), transfer (X3), college biology (X4), college chemistry (X5), college physics (X6), college environmental science (X7), college earth science (X8), part-time (X9), credits taking (X10), first-generation (X11), relative with degree in education (X12), and current GPA (X13). The models used the independent variables or factors in the following combinations:

- Full Model: *X1, X2, X3, X4, X5, X6, X7, X8, X9, X10, X11, X12, X13*
- Academics Model: *X1, X10, X13*
- Uncontrolled Factors Model: *X2, X3, X9, X11, X12*
- Forward Regression: all the independent variables were included in the initial stage of building the model.

Statistical Analyses of Study

The analyses that this study employed included two different statistical approaches. The multiple linear regression and the binomial logistic regression are described below.

Multiple linear regression (MLR) is a type of analysis performed to investigate the extent to which independent variables explain the variance of a dependent variable. The analysis yielded a coefficient of determination (R^2) which explains how well the predictor or independent variables explain the variance in the dependent variable (Courville & Thompson, 2001; Schneider et al., 2010). The closer the value of R^2 is to 1 the stronger the ability of the regression model to explain the variance in the dependent variable (Huang and Fang, 2013).

Binomial logistic regression (BLR) is a type of predictive modeling analysis that can be used to assess the association between a dependent variable and two or more independent variables (Kuha & Mills, 2017). The assessment of the fitness of the logistic model occurs at two levels. The first assessment of the Full model is via the chi-square likelihood ratio test where the null model with no predictors is compared with the full model including all predictors. The second assessment is via the Hosmer & Lemeshow (2000) goodness-of-fit test. The entire model fit is assessed by Nagelkerke's R^2, considered a pseudo R square, and is an explanation of the amount of variation in the dependent variable and the ability of the model to correctly classify preservice teachers group membership of the dependent variable (Smith & McKenna, 2013). The null hypothesis of the model is that βs is equal to "0" and the alternative hypothesis is that at least one β is not equal to "0" (Chao-Ying et al., 2002). The difference between the MLR and the BLR models is that with the BLR model, the dependent variable must be dichotomous or binary such as pass or fail.

While logistic BLR model offers a precise cutoff of preservice teachers' ability to pass or fail the science core of the practice EC-6 certification examination, MLR reveals the direct contribution of each of the independent variables and offers the ability to make an approximation of the actual score earned (Kuha & Mills, 2017).

Data Treatment

For multiple linear regression analyses, the dependent variable, which was the post-test score preservice teachers made on the practice exam, was represented as a continuous variable with a range from 1-100. For binomial logistic regression analyses the scores that preservice teachers obtained on their practice exam were converted from the continuous variable score of ≥ 80 as "pass" which was coded to "1", and $< 80\%$ as "fail" and coded to "0" for the dichotomous result. Some of the data from the survey were converted from "yes" and "no" into dichotomous values "1" and "0" all which are summarized in Table 2.

Each of the independent variables were individually used in a simple linear regression analysis with the dependent variable (the practice exam score), so that one could use a single variable which was statistically significant at $p \leq 0.05$ to predict the performance of preservice teachers who may not have data for all 13 independent variables.

Table 2

Summary of Variable Coding used in the Study

Variables	Description	Variables	Description
X1	Categorical: Pass "1" and Fail "0"	X7	Dichotomous
X2	Categorical		0=Did not take college environmental science
	1=Freshman		1= Took college environmental science
	2=Sophomore	X8	Dichotomous
	3=Junior		0=Did not take college earth science
	4=Senior		1= Took college earth science
X3	Dichotomous	X9	Dichotomous
	0= Did not transfer		0= not part time
	1= Transferred		1=Part-time
X4	Dichotomous	X10	Continuous variable ranges from "0" to "24" or above
	0=Did not take college biology	X11	Dichotomous
	1= Took college biology		0= Not first-generation college student
X5	Dichotomous		1=First-generation college student
	0= Did not take college chemistry	X12	Dichotomous
	1=Took college chemistry		0=No relative with degree in education
X6	Dichotomous		1=relative with degree in education
	0=Did not take college physics	X13	Continuous variable ranges from 2.0 to 4.0
	1=took college physics		

Data Analysis

Both descriptive and inferential statistics were used to assess the data in terms of the numbers and percentages of preservice teachers who: passed the pretest, failed the pretest, passed both the pretest and post-test, and the differences in percentage between the pretest and the post-test. To answer the research question using multiple linear regression analyses, the effects of all 13 independent variables (X1-X13) on the ability of preservice teachers to pass the practice EC-6 TExES certification exam (dependent variable) were analyzed. This was repeated for a three-predictors model (Academics (X1, X10, and X13), a five predictors model (Uncontrolled Factors (X2, X3, X9, X11, X12) and a forward linear regression model.

To answer the research question using the binomial regression analyses, a 13-predictors logistic regression model (Full) was fitted to the data to examine the study's hypothesis of the likelihood that a preservice teacher will pass the practice EC-6 TExES certification exam based on the predictors. This was repeated for a three-predictors model (Academics), a five predictors model (Uncontrolled Factors) and a forward logistic regression model. The data was screened to verify that assumptions have been met.

Results

The impact of the BIOL 1082 course and its effect on the improvement in proficiency on each of the EC-6 TExES Domain IV 18 competencies was conducted by calculating the difference between performance of preservice teachers on pretest and post-test of the practice exam. The 18 competencies were categorized based on what was or will be covered in each of the prerequisite science courses that preservice teachers needed to take prior to taking the official EC-6 TExES examination. Though some of the Domain IV EC-6 TExES competencies were not designated to be covered in the BIOL 1082 course, the gains of all 18 competencies were calculated for the competencies covered in physics, environmental science, earth science, and scientific methods. The distribution and performance on each of the competencies is shown in Table 3. The pre/post-test included 45 questions (Table 1). As shown in Table 3, the increase in performance average (points) for biology was 5.82 points. In descending order, the performance average increase for environmental science was 5.70 points (5 questions); followed by Earth science (2.75; 11 questions), physics (2.64; 8 questions), chemistry (1.28; 5 questions) and science methods (1.5; 17 questions) (See Table 1 for # of questions/competency).

Table 3

Gains in Competencies 1-18 of Domain IV of EC-6 TExES Practice Exam

Competencies Domain IV Science EC-6 TExES	Course at university	Subject matter covered	Pretest average	Posttest average	Change in points
C1, C2, C4, C10, C11, C12, C13, C14	BIOL 1082	Biology	71.29	77.11	+5.82
C8, C9		Chemistry	75.58	76.86	+1.28
C7, C8, C9	PHYS 1210	Physics	66.16	68.80	+2.64
C14, C17	BIOL 1132	Environmental Science	74.88	80.58	+5.70
C15, C16, C17, C18	GEOG 1710	Earth Science	63.68	66.43	+2.75
C3, C5, C6	EDEE 4330	Science Methods	86.46	88.04	+1.58

Descriptive Statistics

Generally, the study showed an improvement in performance on the post-test in comparison to the pretest (Table 4). Independent variables BIOL 1082 (X1), transfer (X3), part-time (X9), first-generation (X11), and relative with degree in education (X12) each had 42 out of 170 preservice teachers who passed the pretest, and 67 out of 170 who passed the post-test. The variable that explained the most variance in the model to predict performance on the practice EC-6 TExES exam was X1 (grade in BIOL 1082). Preservice teachers who failed BIOL 1082 experienced no success on either the pretest or post-test of the practice EC-6 TExES exam.

The overall post-test passing rate was approximately 40%. Freshmen, sophomore, and junior preservice teachers had pass rates between 23% and 34%. Seniors, on the other hand, experienced zero success on the pretest and about 15% success on the post-test. Preservice teachers who were non-transfer students had a 12% higher passing rate in the pretest and 16% higher on the post-test compared to transfer students. Full-time preservice teachers outperformed those who were part-time, students who were not first-generation college students outperformed first generation college students, and preservice teachers who had family members with degrees in education outperformed those who did not.

Compared to preservice teachers taking 15 or fewer credits, preservice teachers taking 16-22 credits had less success in the pretest but then were within the range of performance percentage between 40-49% on the post-test. 55.1% of preservice teachers who had a GPA between 3.51-4.0 passed the post-test. Surprisingly, those with GPAs between 2.0-2.5 followed this with a pass rate of 50%.

Table 4

Descriptive Statistics of Variables

Variables	Categories	# of participants	Passed pretest	%	Passed posttest	%	Passed both	%	Difference Pre/post (%)
Grade in BIOL 1082 (X1)	Passed BIOL 1082	159	42	26.40	67	42.10	36	22.60	15.70
	Failed BIOL 1082	9	0	0	0	0	0	0	0
	Total	168	42	25	67	40	36	21.40	14.70
Classification (X2)	Freshman	37	9	24.30	15	40.50	9	24.30	16.20
	Sophomore	56	18	33.90	25	44.60	14	25	10.70
	Junior	64	15	23.40	25	39.10	13	20.30	15.70
	Senior	13	0	0	2	15.40	0	0	15.40
	Total	170	42	25.30	67	39.40	36	21.20	14.10
Transfer (X3)	Transfer	86	16	18.60	27	31.40	13	15.10	12.80
	Non-transfer	84	26	30.90	40	47.60	23	27.40	16.70
	Total	170	42	24.70	67	39.40	36	21.20	14.70
College Science Courses (X4-X8)	Biology	49	10	20.40	13	26.50	6	12.20	6.10
	Chemistry	14	3	21.40	2	14.30	1	7.10	-7.10
	Physics	52	10	19.20	17	32.70	7	13.50	13.50
	Environmental Science	87	16	18.40	33	37.90	14	16.10	19.50
	Earth Science	66	16	24.20	29	43.90	15	22.70	19.70
Part-time (X9)	Part-time	16	2	12.50	4	25	1	6.25	12.50
	Full-time	154	40	25.90	63	40.90	35	22.70	15
	Total	170	42	24.70	67	39.40	36	21.20	14.70
Credits Taking (X10)	3 to 7	6	0	0	1	16.70	0	0	16.70
	8 to 12	35	11	31.40	17	48.60	8	22.90	17.20
	13 to 15	76	22	28.94	34	44.70	20	26.30	15.80
	16 to 22	25	4	16	10	40	4	16	24
	Total	142	37	26	62	43.70	32	22.50	17.70
First-generation (X11)	Yes	70	12	17.10	21	30	10	14.30	12.90
	No	100	30	30	46	46	26	26	16
	Total	170	42	24.70	67	39.40	36	21.20	14.70

Table 4 (*Continued*)

Variables	Categories	Number of students	Passed pretest	%	Passed posttest	%	Passed both	%	Difference Pre/post (%)
Relative with degree in Education (X12)	Relative	59	18	30.50	30	50.80	16	27.10	20.30
	No relative	111	24	21.60	37	33.30	20	18	11.70
	Total	170	42	24.70	67	39.40	36	21.20	14.70
Current GPA (X13)	2.0-2.5	6	3	50	3	50	3	50	0
	2.51-3.0	16	1	6.25	7	43.70	0	0	37.50
	3.01-3.5	42	8	19	9	21.4	5	11.90	2.40
	3.51-4.0	49	18	36.70	27	55.10	17	34.70	18.40
	Total	113	30	26.50	46	40.70	25	22.10	14.20

Multiple Linear Regression Univariate Analyses

Studies have suggested that in a multiple linear regression there should be at least 10 observations per independent variable (Sperandei, 2014). Shown in Table 5 are independent variables X1-X13 used in univariate or simple linear regression analyses with the dependent variable. As shown in Table 5, X1, X2, X3, X4, X10, and X13 were statistically significant at $p< 0.05$. The coefficient of determination (R^2), which is the proportion of variance in the practice test score that was explained by the independent variable, are explained in terms of: BIOL 1082 (X1), which explains 16.3% of the variance in the practice exam score, followed by current GPA (X13), then number of credits taking (X10), and transfer status (X3), which each individually explained approximately 5% of the variance in the dependent variable. College biology (X4) follows at approximately 3.1% and then classification (X2) at 2.4%. Independent variables college chemistry (X5), college physics (X6), college environmental. Science (X7), college earth science (X8), part-time (X9), first-generation (X11), and relative with degree in education (X12) each explains <2% of the variance in the dependent variable. Though most studies suggest pursuing variables with a p-value of 0.05 or less, Hosmer and Lemeshow (2000) recommend that a p-value of 0.25 or less can be pursued to avoid the loss of variables that may be valuable to the research, or variables that when combined with another variable may present compelling evidence of effect on dependent variable. Variables X5, X6, X7, and X12, which had levels of significance above 0.25, were kept in the Full model since previous studies have suggested that prior knowledge in subject areas as well as having a relative with similar experience may influence performance on student achievement and thus on the practice EC-6 TExES examination (Kim & Corcoran, 2018).

Table 5

Output of each of the Independent Variables Individually Regressed with the Dependent Variable Performance on Practice Exam

Variable	R	R^2	Adjusted R squared	Sig.	B_1	Standard Coefficient Beta
X1	0.404	0.163	0.158	0*	0.471	0.404
X2	0.154	0.024	0.018	0.045*	-1.71	-0.154
X3	0.219	0.050	0.042	0.004*	-4.352	-0.219
X4	0.176	0.031	0.026	0.022*	-3.860	-0.176
X5	0.093	0.009	0.003	0.228	-3.361	-0.093
X6	0.080	0.006	0.001	0.302	-1.716	-0.080
X7	0.072	0.005	-0.001	0.354	-1.383	-0.072
X8	0.130	0.017	0.011	0.091	2.647	0.130
X9	0.120	0.014	0.008	0.120	-4.075	-0.120
X10	0.224	0.050	0.043	0.007*	0.741	0.224
X11	0.107	0.011	0.006	0.165	-2.159	-0.107
X12	0.083	0.007	0.001	0.286	1.718	0.083
X13	0.236	0.056	0.047	0.012*	4.630	0.236

*Indicates statistically significant at $p \leq 0.05$

Table 6

Model Summary of the Multiple Linear Regression Prediction Models

Name of Multiple Linear Regression Models (MLR)	R^2	p	β_0	β_xX1	β_xX2	β_xX3	β_xX4	β_xX5	β_xX6	βxX7	βxX8	βxX9	βxX10	βxX11	βxX12	βxX13
Full MLR	0.323	0.000*	24.426	0.548	-1.743	-0.795	-1.530	-4.683	0.224	2.726	1.637	3.626	0.807	-0.990	2.609	-1.013
Academics MLR	0.286	0.000*	17.877	0.486									0.726			2.013
Uncontrolled MLR	0.039	0.042*	78.629		-0.258	-4.025						-1.916				
Forward MLR 1	0.274	0.000*	21.504	0.647												
Forward MLR 2	0.313	0.000*	18.332	0.562									0.747			

*Indicates statistically significant at $p \leq 0.05$

Multiple Linear Regression Model Summary

A multiple linear regression model was fitted to the data to evaluate the relationship between preservice teachers score on the practice exam and all 13 of the independent variables in the Full model summarized in Table 6. The model was statistically significant (F (13,93) = 4.451, p< 0.05) with an adjusted R^2 value of 0.323 interpreted as 32.3% of the variance in the practice exam score can be explained by the model.

The model summary for the Academics model included the variables BIOL 1082 (X1), number of credits taking (X10), and current GPA (X13). The model was found to be statistically significant (F (3, 94) = 13.977, p< 0.05), with an adjusted R^2 of 0.286 which is interpreted as 28.6% of the variance in practice exam score can be explained by the Academics Model as shown in Table 6.

The Uncontrolled Factors model included the variables classification (X2), transfer status (X3), part time (X9), first generation (X11), and relative with a degree in education (X12). The model was found to be statistically significant F (5, 161) = 2.365, p< 0.05), with an R^2 of 0.039, which is translated as 3.9% of the variance in the practice exam is explained by the Uncontrolled Factors Model as shown in Table 6.

In Forward linear regression Model 1 was statistically significant (F (1, 93) = 36.421, p< 0.05 with an R^2 value of 0.274 which explains that 27.4% of the variance in practice exam score can be explained by be explained by BIOL 1082 (X1). In Model 2, (F (1, 92) = 6.283, p< 0.05 with an R^2 value of 0.313 which explains that 31.3% of the variance in practice exam score can be explained by be explained by BIOL 1082 (X1) and number of credits taking (X10) as shown in Table 6.

Multiple Linear Regression Coefficients Summary

Coefficients that describe the relationship between the independent variable and the dependent variable provide information about the amount of increase or decrease in a practice exam score that can be predicted by a single unit increase in the independent variable. The coefficients of each of the four multiple regression models are summarized in Table 6.

In the Full MLR model (Table 6), the independent variable, BIOL 1082(X1), was shown to be statistically significant at p<0.05. The coefficient for X1 is 0.548 which is interpreted as for every unit increase in a student's Biology grade in the BIOL 1082 course, a 0.548-unit increase is predicted in the practice exam score, holding constant all the other independent variables. In this model, a preservice teacher is anticipated to have an increased chance of passing the practice examination if they passed BIOL 1082; a decreased chance of passing based on their seniority in classification (X2) as they progress from freshman to senior; a decreased chance of passing the practice exam if they are a transfer student (X3); a decreased chance of passing the practice exam if they took a college biology or chemistry course (X4 and X5); an increased chance of passing if they have taken college physics, environmental science, or earth science (X6, X7, and X8); an increased chance of passing for preservice teachers who identified as part-time (X9) and if they are have a relative with a degree in education (X12); an increased chance of passing the practice certification exam as the number of credits they are taking, (X10) increase in quantity; and a decreased chance of passing if they are first-generation college students (X11), and for every unit increase in current GPA (X13).

In the Academics MLR model (Table 6), both BIOL 1082 (X1), and number of credits taking (X10) were statistically significant at p<0.05. In this model, the coefficient for BIOL 1082 (X1), 0.486 is interpreted as for every unit increase in BIOL 1082 (X1), the predicted score on the practice exam is expected to increase by 0.486 points. For both number of credits taking (X10) and current GPA (X13) for every unit increase in these two variables, the predicted score on the practice exam is predicted to increase by 0.726 points and by 2.013 points respectively.

In the Uncontrolled Factors MLR model (Table 6), independent variable transfer status (X3) was statistically significant at $p<0.05$. For independent variables classification (X2), part-time (X9), and first-generation (X11), a decrease is predicted in the practice exam score for these variables. For transfer status (X3) , the practice exam score is predicted to be 4.0253 points lower for preservice teachers who transferred compared to those who are not transfer students. For relative with degree in education (X12) the practice exam score is predicted to be 1.345 points higher for preservice teachers who have a relative with a degree in education compared to those who did not.

Finally, for the forward MLR model (Table 6), independent variable BIOL 1082 (X1) was placed into the model first and it was statistically significant at $p<0.05$. The second independent variable that was pulled into the forward regression model was number of credits taking (X10). Both were statistically significant and an increase in the practice exam score was predicted for these two variables.

The Full multiple linear regression model, the Academics multiple linear regression model and the Forward linear regression model yielded values of R^2 which explained between 28%-32 % of the variance in the dependent variable. The variables BIOL 1082 (X1), transfer (X3), number of credits taking (X10) and current GPA (X13) were the only variables that maintained statistical significance at $p<0.05$ in one or more of the multiple linear regression models. The Uncontrolled Factors multiple linear regression model did not yield a model that can explain a high enough variance in the dependent variable at just 3.9%. However, the null hypothesis was rejected in knowledge that some of the independent variables do show evidence of impacting preservice teachers score on the practice EC-6 examination.

Binomial Logistic Regression Univariate Analyses Summary

In Table 7, the dependent variable was independently regressed upon each of the independent variable in a univariate logistic regression analysis of which Bio Grade (X1), Transfer (X3), Bio Course (X4), Chem Course (X5), First-generation (X11), Relative with degree in Education (X12), and Current GPA (X13) were statistically significant at $p \leq 0.05$ Hosmer and Lemeshow (2000) recommended that independent variables that are not statistically significant may be kept in a prediction model due to the impact they may have on the dependent variable and may experience increased significance or decreased significance when in conjunction with other independent variables. While some of the other variables did not qualify for Hosmer and Lemeshow's (2000) argument of keeping variables whose *p*-value was ≤ 0.25, the independent variables college environmental science (X7), college earth science (X8), and number of credits aking (X10) were kept due to supporting evidence from other studies that these variables do influence student performance on examinations (Kim & Corcoran, 2018).

Binomial Logistic Regression Variables in the Equation

The column for Y intercept "B" is the coefficient of the equation and describes the relationship between the independent variable and the dependent variable which informs about the amount of increase or decrease in log odds for passing the practice exam that can be predicted by a single unit increase in the independent variable. The results are summarized for the four different logistic models and presented in Table 8.

In the Full Logistic Model, shown in Table 8, χ^2 of 33.265 (13, N= 97, $p < .05$) only the independent variables grade in BIOL 1082 (X1) and Relative with degree in Education (X12) were shown to be statistically significant at $p < 0.05$. In this model, the log of the odds that a preservice teacher passes the practice EC-6 TExES certification exam was positively related to their grade in BIOL 1082. The coefficient is 0.096 and the ODDS ratio is 1.101 which is interpreted as for every unit increase in BIOL 1082 (X1), the logit or the odds of passing the practice exam increases by 0.096-unit and that exponentially for ODDS ratio translates to the odds of passing the practice exam increasing by 1.101

(Table 7). Classification (X2), transfer status (X3), college biology or college chemistry (X4 and X5), part-time status (X9), first-generation status (X11), and current GPA (X13), each with an ODDS ratio <1, are associated with decreased odds of passing the practice EC-6 TExES certification exam for every unit increase in that variable. The variables College physics (X6), college environmental science (X7), or college earth science (X8), , and Relative with degree in Education (X12) each with a positive β and an ODDS ratio >1, are associated with increased odds of passing the practice EC-6 TExES certification exam for every unit increase in that variable.

Table 7

Independent Variable Individually Regressed with the Dependent Variable in Binomial Logistic Regression

Independent Variable	Chi-square *df* (1)	*p-value*
X1 (exam score)	12.418	0*
X2 (classification)	1.46	0.227
X3 (transfer status)	4.707	0.03*
X4 (college biology course)	4.95	0.026*
X5 (college chemistry course)	4.597	0.032*
X6 (college physics course)	1.437	0.231
X7 (environmental science course)	0.164	0.686
X8 (Earth science course)	0.923	0.337
X9 (employed part-time	1.624	0.203
X10 (# of credits taken)	0.593	0.441
X11 (1st generation student)	4.479	0.034*
X12 (relative in education)	4.717	0.03*
X13 (GPA)	4.297	0.038*

Note. *Indicates statistically significant at $p \leq 0.05$

Per the Academics Logistic Regression Model, as shown in Table 8, χ^2 of 12.834 (3, N= 97, $p<.05$) only grade in BIOL 1082 (X1) is statistically significant at $p<0.05$. For X1, the coefficient is 0.067 and the ODDS ratio is 1.069, which is interpreted as for every unit increase in X1, the logit or the odds of passing the practice exam increases by 0.067-unit and that exponentially for ODDS ratio translate the odds of passing the practice exam increased by 1.069. For independent variables credits taking (X10) and current GPA (X13) though not statistically significant in the model, both are associated with an increase in the odds of passing the practice EC-6 TExES certification exam.

For the Uncontrolled Factors Logistic Model, as shown in Table 8, χ^2 of 13.119 (5, N= 169, $p<.05$) transfer status (X3) was statistically significant at p<0.05. Independent variables classification (X2), transfer (X3), part-time (X9), and first-generation (X11) are each associated with a decrease in the ODDS chance of passing the practice exam with an odds ratio <1. For the independent variable relative with degree in education (X12), the odds of passing the practice exam increases by 0.597 unit and exponentially for the ODDS ratio the odds of passing the practice exam is 1.816 times more likely for a preservice teacher who has a relative with a degree in education compared to those who do not.

In the forward logistic regression model, as shown in Table 8, grade in BIOL 1082 (X1) was pulled into the model first yielding the results summary χ^2 of 11.042 (1, N= 97, $p<.05$) with the odds of passing the practice exam increasing by 0.093-unit and exponentially for ODDS ratio this translates to the odds of passing the practice exam increasing by 1.098. In step 2 of the model, relative with a degree in education (X12) was pulled in addition to BIOL 1082 X1. Both X1 and X12 were statistically significant and their positive β coefficient and odds ratio values <1 were associated with an increased likelihood and ODDs of preservice teachers passing the science portion of the practice EC-6 TExES certification examination which can be seen in Table 8.

Table 8

Coefficient of Variables in the Equation and Goodness of Fit for Logistic Models

Binomial Logistic Regression Model	Predictors	β	*SE* β	Wald's χ^2	*df*	*p*	e^{β} (odds ratio)	95% C.I.for EXP(B) Lower	Upper
Full Logistic Regression Model	Constant	-9.674	3.998	5.856	1	0.016*	NA		
	X1 exam score	0.096	0.039	5.983	1	0.014	1.101	1.019	1.19
	X2 classification	-0.263	0.369	0.506	1	0.477	0.769	0.373	1.585
	X3 transfer status	-0.271	0.727	0.138	1	0.71	0.763	0.183	3.174
	X4 college bio course	-0.687	0.75	0.841	1	0.359	0.503	0.116	2.185
	X5 college chem course	-21.112	14907	0	1	0.999	0	0	0
	X6 college phyics course	0.565	0.623	0.822	1	0.365	1.759	0.519	5.962
	X7 college env sci course	0.558	0.584	0.914	1	0.339	1.748	0.556	5.493
	X8 college Earth sci	0.207	0.553	0.14	1	0.708	1.23	0.416	3.633
	X9 work part-time	-0.047	1.176	0.002	1	0.968	0.954	0.095	9.56
	X10 # credits	0.099	0.12	0.679	1	0.41	1.104	0.873	1.397
	X11 1st generation	-0.224	0.549	0.166	1	0.684	0.8	0.272	2.347
	X12 relative in edu	1.56	0.6	6.75	1	0.009*	4.759	1.467	15.438
	X13 GPA	-0.071	0.578	0.015	1	0.902	0.932	0.3	2.893
	Overall model evaluation			χ^2	*df*	*p*			
	Goodness-of-fit test Hosmer & Lemeshow			10.06	8	0.261			
Academic Logistic Regression Model	Constant	-9.264	2.872	10.407	1	0.001	0		
	X1 exam score	0.067	0.033	4.222	1	0.04*	1.069	1.003	1.14
	X10 # credits	0.092	0.082	1.261	1	0.262	1.097	0.934	1.288
	X13 GPA	0.544	0.478	1.295	1	0.255	1.724	0.675	4.403
	Overall model evaluation			$\chi2$	*df*	*p*			
	Goodness-of-fit test Hosmer & Lemeshow			21.53	8	0.006			
Uncontrolled Factors Logistic Regression Model	Constant	-0.223	0.494	0.204	1	0.651			
	X2 classification	0.119	0.234	0.259	1	0.611	1.127	0.712	1.784
	X3 transfer status	-0.88	0.43	4.188	1	0.041*	0.415	0.178	0.963
	X9 work part-time	-0.346	0.643	0.289	1	0.591	0.708	0.201	2.496
	X11 1st generation	-0.591	0.37	2.553	1	0.11	0.554	0.268	1.143
	X12 relative in edu	0.597	0.362	2.711	1	0.1	1.816	0.893	3.694
	Overall model evaluation			$\chi2$	*df*	*p*			
	Goodness-of-fit test Hosmer & Lemeshow			10.79	8	0.214			

Table 8(*Continued*)

	Binomial Logistic Regression Model	Predictors	β	*SE* β	Wald's χ^2	*df*	*p*	e^{β}(odds ratio)	95% C.I.for EXP(B) Lower	Upper
Step 1		X1 exam score	0.093	0.031	9.009	1	0.003*	1.098	1.073	1.269
		Constant	-8.274	2.681	9.525	1	0.002	0		
Step 2		X1 exam score	0.098	0.031	9.88	1	0.002*	1.103	1.078	1.275
		X12 relative in edu	1.343	0.488	7.587	1	0.006*	3.83	1.324	13.448
		Constant	-9.167	2.728	11.29	1	0.001	0		
Step 3		X1 exam score	0.11	0.033	11	1	0.001*	1.116	1.093	1.318
	Forward Logistic Regression Model	X5 college chem course	-21.589	14533	0	1	0.999	0	0.000	
		X12 relative in edu	1.555	0.526	8.721	1	0.003*	4.733	1.688	23.759
		Constant	-10.094	2.896	12.15	1	0	0		
		Overall model evaluation			χ^2	*df*	*p*			
Step 1		Goodness-of-fit test Hosmer & Lemeshow			6.654	8	0.574			
Step 2		Goodness-of-fit test Hosmer & Lemeshow			6.578	8	0.583			
Step 3		Goodness-of-fit test Hosmer & Lemeshow			4.39	8	0.82			

*Indicates statistically significant at $p \leq 0.05$.

Binomial Logistic Regression Prediction Accuracy

In Table 9, the classification table of the model with all 13 variables (Full Logistic Model) has an overall prediction accuracy of 75.26%. In the "failed" practice test row, a total of 56 (45 +11) were observed as failures. However, the model correctly predicted 45 out of the 56 (80.4%) of those to be failures and incorrectly predicted 11 out of 56 (19.6%) as passes. In the "passed" practice test row out of the 41 observed to have passed, the model incorrectly predicted 13 out of 41 (31.7%) as failures and 28 out of 41 (68.3%) as passes.

The Academics Logistic Model in Table 9 had an overall prediction accuracy of 66.3%. In the "failed" practice test row, a total of 57 (45 +12) were observed as failures. However, the model correctly predicted 45 out of the 57 (78.9%) of those to be failures and incorrectly predicted 12 out of 56 (21.1%) as passes. In the "passed" practice test row out of the 41 observed to have passed, the model incorrectly predicted 21 out of 41 (51.2%) as failures and 20 out of 41 (48.8%) as passes.

The prediction accuracy for the Uncontrolled Factors Logistic Model in Table 9 was 62.3% for correctly predicting 83.3% of preservice teachers who failed the practice test and 30.8% accuracy for correctly predicting preservice teachers who passed the practice test.

Forward logistic regression prediction accuracy in Table 9: Model 1 had a prediction accuracy of 65.9%, Model 2 had an overall prediction accuracy of 70.1% and the overall prediction accuracy for Model 3 was 72%. Model 1, 2, and 3 all had a prediction accuracy for preservice teachers who failed the practice at 79%. The prediction accuracy for preservice teachers who passed the practice test for all three forward logistic models were: Model 1 (49%), Model 2 (59%), and Model 3 (63%). The probability of correctly predicting the correct group membership increases (sensitivity) as the probability of predicting the incorrect group membership decreases (specificity).

Table 9

Classification Table of the Prediction Accuracy for Logistic Regression Models

	Logistic Regression Model	**Observed**	**Predicted**		**% Correct**
			Failed practice test	Passed practice test	
	Full	Failed practice test	45	11	(45/56) 80.40%
		Passed practice test	13	28	(28/41) 68.30%
		Overall % correct			(73/97) 75.26%
	Academics	Failed practice test	45	12	(45/57) 78.9%
		Passed practice test	21	20	(20/41) 48.8%
		Overall % correct			(65/98) 66.3%
	Uncontrolled Factors	Failed practice test	85	17	(85/102) 83.3%
		Passed practice test	47	21	(21/68) 30.8%
		Overall % correct			(106/170) 62.3%
Step 1	Forward Logistic Regression Model	Failed practice test	44	12	(44/56) 78.6%
		Passed practice test	21	20	(20/41) 48.7%
		Overall Percentage			(64/97) 65.9%
Step 2		Failed practice test	44	12	(44/56) 78.5%
		Passed practice test	17	24	(24/41) 58.5%
		Overall Percentage			(68/97) 70.1%
Step 3		Failed practice test	44	12	(44/56) 78.5%
		Passed practice test	15	26	(26/41) 63.4%
		Overall Percentage			(70/97) 72.2%

Discussion

In some longitudinal and standalone studies conducted to gauge students' success, there is supporting evidence suggesting that academic readiness and success on examinations can be predicted by GPA, full-time status, high self-efficacy, coaching, and other environmental factors (Frizzell, 2014; Gard, 2011; Kazempour & Sadler, 2015; Kim & Corcoran, 2018). While most of these studies focused on factors that may impact the probability of passing a course and others focused on passing a standardized test, most of these studies clumped different subject matters together, neglecting the fact that there are some factors that affect performance that may be unique to specific subject matters (Bains, 2011; Warren 2017).

The *Every Student Succeeds Act* (ESSA) brought changes to the scoring and number attempts that preservice teacher have when taking their state certification examination as well as it increased the standards for Educator Preparation Programs (EPPs) (Every Student Succeeds Act, 2015). EPPs while tasked with preparing preservice teachers for their state certification examination, are also faced with maintaining their accreditation status at 85% pass rate in one academic year (Rickenbrode et al. 2018; Warren, 2017). In Texas, preparation for the Texas Examinations of Educator Standards (TExES) certification examination by (Early Childhood to Six) EC-6 preservice teachers include limited attempts at passing their state certification examination at only five attempts as well as more stringent scoring whereby they are now faced with earning an 80% on each of the five subject matter which means proficiency of each subject matter is mandatory (TEA, 2016). Prior to ESSA 2015, EC-6 TExES preservice teachers were able to mask their inadequacies of a specific subject matter because the scores of each of the five core subjects were combined into one score and if that score was at or above 80%, it was considered a pass (TExES™ program preparation manual, 2018; Warren, 2017). It is now crucial that EC-6 TExES preservice teachers obtain mastery in each of the five subject matters of: English Language Arts Reading, the Science of Teaching Reading, Mathematics, Social Studies, Science, and Fine Arts, Health, and Physical Education (TExES™ program preparation manual, 2018).

Bains (2011) investigated preservice teachers' pass rate of individual subject matters of EC-6 Generalist examination, but the study did not examine factors that can impact actual scores. Gard (2011) examined the prediction of performance on the individual subject areas, such as preservice teachers' performance on TExES 8-12 history, which is a standalone certification subject matter and does not have other variables which may impact performance based on different subject matters. Fang and Wang (2013) investigated the best statistical tools to predict performance at the best performance accuracy, but the focus was on how well an engineering course performance can be predicted using different models. Warren (2017) investigated predictive factors that can influence teacher candidates in EPPs, but this study focused on competing factors that can impact preservice teachers before they even start the EPP program. Another study focused on comparing performance on the content examination vs the pedagogy part for early childhood preservice teachers (Capraro et al., 2005). Other studies focused on preservice teachers' perceptions and self-efficacy (Kim & Corcoran, 2018). Another study (Corcoran and O' Flaherty, 2018) focused on factors that can predict preservice teachers' effectiveness in classroom teaching and thus was not focused on factors impacting the actual exam. These studies either explored performance on the entire certification examination, or on other pre-requisite factors, and thus far, no single study has narrowed down their investigation into the performance on distinct subject matters within a single certification examination. Creating predictive model(s) based on potential independent variables or factors that can predict EC-6 TExES preservice teachers' performance on the individual subject matter of science within the content portion of the certification examination was the focus of this study and serves to fill the gap in the literature which can inform EPPs teachers and preservice teachers alike in every early childhood to six grade certification programs everywhere.

Several studies investigating factors that can influence preservice teachers' success on certification examinations highlighted factors such as age, gender, ethnicity, workload, GPA, familial influence, and performance in their tertiary education program that were significant influencers on their performance on the overall examination (Bain, 2011; Gard, 2011; Huang & Fang, 2013; Warren, 2017). However, most of these studies were focused on holistic examination performance where more than one subject area was combined. Based on different attitudes, perceptions, confidence of proficiency that preservice teachers have on different subject matters this study explored some of the prior factors as well common ex post facto factors that preservice teachers listed which may impact their performance on the science portion of the EC-6 TExES certification examination.

A mixture of descriptive statistics, multiple linear regression, and binomial logistic regression were used in this study to investigate the impacts, if any, that the independent variables had on the practice exam success (Huang & Fang, 2013; Warren, 2017). Researchers preferring multiple linear regression have argued that while binomial logistic regression may successfully predict an individual's ability to pass or fail, a pass may be within boundary of a failure and that person may still be in jeopardy of failing the examination (Huang & Fang, 2013). The employment of both multiple linear regression and binomial logistic regression result in the added benefit that while a binary logistic regression predictive model can help predict a preservice teacher's likelihood of passing or failing the examination due to discrete values associated with either a "pass" or "fail", a multiple linear regression prediction model can offer a more precise score for preservice teachers on a scale from 0 to 100 (Huang & Fang 2013; Khajuria, 2007).

Using MLR to build predictive models has been a strategy used to gauge success in many academic settings. Use of the coefficient of determination (R^2) along with statistical significance have been used as rapid analytic tools for exploring whether predictive variable(s) have impact on the performance outcome of learners (Karamazova et al., 2017; Yang et al., 2018). Each of the MLR models was statistically significant at $p \leq 0.05$ as shown in Table 10. The Full model explained 32.3% of the variance in the practice exam score. The Academics model explained 28.6% of the variance in the practice exam score. The Uncontrolled Factors model explained only 3.9% of the variance in the practice exam score. The forward regression model explained 31.3% of the variance in the practice exam score. The predictor variable which was most often found in all models to be statistically significant was Grade in Biol 1082 (X1).

Table 10

Multiple Linear Regression Models Comparison

Model	R^2	Adjusted R^2	p-value
Full Model (All 13 variables)	0.417	0.323	0.000*
Academic Model	0.308	0.286	0.000*
Uncontrolled Factors Model	0.068	0.039	0.042*
Forward Regression Model	0.327	0.313	0.014*

*Indicates statistically significant at $p \leq 0.05$.

The prediction accuracy of each of the binomial logistic model is summarized in Table 11. Remarkably, the models provided better prediction for the likelihood of failure rather than the likelihood for success on the practice exam. Given that the pass rate on the pretest was 42 out of 170 (24.7%) and the pass rate on the post-test was 67 out of 170 (39.4%), then there may be an advantage of the model predicting failure so that the preservice teachers predicted to fail can be identified and recommended for remediation prior to their official EC-6 TExES examination (Shipe et al., 2019). Each model was able to predict preservice teachers at risk of failing with an accuracy of at least 79%. The full model was the best at predicting who was likely to fail at 80.40%, pass at 68.3% and overall, at 75.30%. The Uncontrolled factors model was the most accurate at predicting failure at 83.3% and simultaneously the least accurate at predicting success at 30.8%. The forward logistic model correctly identified 79% of the preservice teachers expected to fail and 59% expected to pass. The overall prediction accuracy of the models also suggests that the Full logistic model provided the best overall predictor model (75.30%). All four logistic regression model were statistically significant at $p \leq 0.05$ which rejects the null hypothesis that states that the independent variables will not have any effect on preservice teachers' performance (Huang & Fang 2013).

Table 11

Binomial Logistic Regression Models Comparison

Prediction Model	Correctly identified for failing practice test (At-risk))	Correctly identified for passing practice test	Overall % Correct
Full Model (all 13 variables)	80.4% (45 out of 56)	68.3% (28 out of 41)	75.30
Academics Logistic Model	78.9% (45 out of 57)	48.8% (20 out of 41)	66.30
Uncontrolled Factors Logistic Model	83.3% (85 out of 102)	30.8% (21 out of 68)	62.30
Forward Regression Logistic Model	79% (44 out of 56)	59% (24 out of 41)	70

Based on the findings of the study, the full predictive models for the MLR model explained the largest variance in the dependent variable at 32.3% and for the BLR model correctly predicted an overall 75.26% of preservice teachers' successes and failures on the practice exam. However, the full model consists of 13 independent variables, and it may not always be possible to collect more than ten data points for each preservice teacher to successfully use the full model. For this reason, other models may better serve the purpose to predict success on the certification examination. In this study, the simplest model that was most significant in predicting performance on the practice exam was the grade in BIOL 1082 (X1). This model may easily be used for an EPP to evaluate and predict a whole class of preservice teachers' ability to succeed on the science domain IV EC-6 TExES certification exam (Huang & Fang, 2013) Both forward MLR and BLR predictive models also included X1 as the first independent variable in the model. The Academics MLR and BLR model included grade in BIOL 1082 (X1), credits taking (X10), and current GPA (X13) and may also be useful in predicting performance in being the next best model with the smallest subset of independent variables. Uncontrolled Factors MLR and BLR models were not very good at predicting success, instead however the variable transfer (X3) within the model was a good predictor variable that could be used to identify individuals with transfer status as those which may be in the greatest need of remediation and intervention to help them prepare for their official EC-6 TExES certification examination.

The research findings gave rise to some important conclusions: (1) The best predictor of performance on the practice EC-6 TExES certification examination was the grade earned in BIOL 1082, Biology for Elementary Educators. (2) The independent variables transfer status (X3), Credits Taking (X10), and Current GPA (X13) individually also predicted performance and maintained statistical significance (3) The prediction accuracy for predicting passing the practice examination was between 30%-50% and predicting failure was at and above 78%. In both the MLR and the BLR predictive models, which included all 13 independent variables, explained the most variation and had the highest prediction accuracy in terms of performance on the practice examination. In the case of this study, it was important to predict which EC-6 TExES preservice teachers were in danger of failing to stage interventions prior to their official certification examination at this university. However, we posit that these same tools can be used by other EPPs to predict their own student's preparedness for certification success.

Other EPPs could use the individual predictor model (X1) to predict performance of preservice teachers at the end of a semester when students have taken any science for educators' course. Additionally, a single preservice teacher's performance can be predicted using the Full MLR model and the Full BLR models because both MLR model explained the largest variation in the dependent variable and the BLR model yielded the largest prediction accuracy (Huang & Fang 2013). If the EPP does not have data for all 13 variables, the academics model can still predict performance with only three independent variables, grade in Science for Educators course (X1), credits taking (X10), and current GPA (X13). This model is the next best predictor of performance on the practice EC-6 TExES certification examination with the least number of variables.

Conclusions

Pursuing any type of degree or certification in tertiary education can be met with some type of financial pressure (Feuer et al., 2013). Additionally, in repeatedly failing their certification examination, preservice teachers can decrease the likelihood of starting a teaching career in a reasonable amount of time and may also result in the diminishing of confidences and the lessening of self-efficacy (Masters, 2018). Prediction of preservice teachers' performance on their certification examinations have been explored by other researchers (Hutson, 2018; Warren, 2017). However, no other study thus far looked at predicting the performance by preservice teachers on the science portion of their official EC-6 TExES certification examination. And most importantly, no other study has created predictive models for determining preservice teacher's success on the science portion of their EC-6 TExES certification exam.

In this study, variables that can possibly influence preservice teachers' performance on the science core of their EC-6 TExES certification examination were examined. Voluntary participants of this study were preservice teachers enrolled in BIOL 1082, is a mandatory science course EC-6 preservice teacher need to take in preparation for their official EC-6 TExES exam. This course covered almost half of the competencies preservice teachers need to be proficient in to successfully pass the official EC-6 TExES certification examination (TExES™ program preparation manual, 2019). The BIOL 1082 course included "clicker" questions over all the 18 competencies of Domain IV EC-6 TExES certification exam which were woven within each lesson taught to preservice teachers. Numerous hands-on lessons along with "think-pair-share" activities among others allowed for numerous opportunities for preservice teachers to learn concepts as well as conduct deep discussions on key topics. The practice exam was issued in the beginning of the semester of BIOL 1082 (pretest) and at the end of the semester (post-test). The Qualtrics™ survey was done online. The independent variables in this study were part of the survey which collected ex post facto, qualitative, and quantitative data from preservice teachers and the post-test score on the practice exam was the dependent variable of the study.

The creation of predictive models was conducted by use of multiple linear regression (MLR) and binomial logistic regression (BLR). The R^2, statistical significance and regression coefficient of each independent variable within each model were examined for the MLR models. For the BLR models the Nagelkerke R^2, the statistical significance, the odds ratio and the classification with prediction accuracy were examined. The results from the descriptive, MLR and BLR models analyses suggests that the Grade in BIOL 1082 (X1) is the most useful independent variable in predicting preservice teachers' performance on the science core of the practice EC-6 TExES certification examination.

The Academics model, which included the variables: grade in BIOL 1082 (X1), credits taking (X10), and current GPA (X13) for both the MLR and BLR predictive models were statistically significant and generally number of credits taking (X10), and current GPA (X13) trended toward an effect and toward increased chances of passing the practice exam. The only variable with statistical significance in the Uncontrolled Factors MLR and BLR models was transfer (X3) and revealed that transfer students' regression coefficient of -4.025 translated to the practice exam score prediction of 4.025 points lower for preservice teachers who transferred compared to those who had not. EPPs with information such as this may seek to offer more scaffolding and remediation to the preservice teachers that transfer into their program.

GPA, which in some studies have been shown to be a reliable predictor of students' performance, in that the GPA the range of 3.51 to 4.0 had the highest percentage pass at 55%. However, this trend did not continue in succession for the other GPA ranges. This correlation between GPA and success on practice exam could be strengthened if EPPs continue to maintain and improve the rigor of their curriculum to match the rigor of the TExES exam. Fine tuning of these efforts can be attained by collaboration between EPPs and TExES department to align curriculum tightly with expectations on the exit examination.

Though some of the predictors were not statistically significant within the model, their effect on the practice exam score were noticeable and, in some instances, had a small effect. Independent variable part-time (X9), revealed that these preservice teachers were less successful compared to those who were not. First-generation (X11) preservice teachers were less successful than their counter parts. The preservice teachers with a relative in education were more successful than those without. Preservice teachers who took college environmental science or college earth science classes (X7 and X8) generally trended with increased score or performance on the practice exam. For preservice teachers who took college physics (X6), while there was no direct increase in performance this variable seemed to act like a suppressor variable whose role is to indirectly impact the practice exam performance. Classification (X2) did not reveal sizeable differences between freshman, sophomores, and juniors' performance on the practice exam. However, preservice teachers classified as seniors who are most likely on the cusp of taking their official EC-6 TExES exam had the lowest performance on the practice exam with 0% passing the pretest and 15.40% of them passing the post-test. As such, the performance for seniors is crucial and should be prioritized. Having taken a college biology or chemistry class (X4 and X5) were associated with decreased performance on the practice exam which was also surprising but unfortunately what was missing from the survey questions and what was not revealed was where the preservice teachers took those courses and whether they had passed them.

Practice examinations have been shown to increase confidence and reduce anxiety of test takers (Bandura, 1997; Gard, 2011; Sullivan et al., 1996). EPPs could use practice examinations as a form of assessment which can provide insight as to preservice teachers' readiness for their certification examination as well as make available early interventions where needed. This practice certification examination may be used by EPPs in all their mandatory courses to find out what gains, if any, are made by preservice teachers and may help canvas the unremitting visibility of students' readiness as they progressively move closer to taking their official examination. This can create transparency that is advantageous to both EPPs and preservice students as they map and document their preparedness for the official TExES exam.

The objective of the predictive models is to allow for well-timed content interventions, when required, by the identification of preservice teachers who may be in danger of failing the certification examination. Such knowledge allows preservice teachers the chance to receive backing and scaffolding of content in practices that increase teacher content knowledge. Both types of regression predictive models suggest EPPs can predict preservice teachers who may be at risk of failing their certification exam. The binomial logistic regression offers the "big picture" of the probability of a preservice teacher passing or failing the exam, and multiple linear regression will give a predicted score that reveals borderline students.

Predictive modeling is and has been commonly used in a variety of pursuits, including retail, healthcare, entertainment, manufacturing, cybersecurity, human resources, sports, politics, and weather for 20 plus years, but is far less commonly used in education. In fact, in education, the most common uses deal with student retention indicators (Al Sheeb, et al. 2019; Bird, et al., 2021; Hung, et al., 2019; Smith, et al., 2012). However, as this study shows, predictive modeling can play an invaluable role in teacher education preparation as well. As many other industries have realized, predictive modeling is a means by which one can tentatively see into the future to determine a potential outcome, and by which to make decisions about resource management. This study demonstrates that predictive modeling is just as effective in the field of teacher education and can serve to provide important information that can be used to aid our perspective student teachers on their journey to their future career.

Limitations

The data used in this study was obtained from the survey which was self-reported. Some factors that may affect preservice teachers' performance on the certification exam may not have been used in this study. Factors such as the use of psychological, emotional, and other abstract variables that may be difficult to quantify may play a crucial role in preservice teachers' ability to pass the practice examination and these variables were not included in the study (Huang & Fang, 2013). Factors such as technology, the amount of time spent on social media, learning methods, self-motivation, and intrinsic reward system, along with some recent factors that might have not yet been studied for their effects on students' performance. Additionally, the time frame between completion of BIOL 1082 and taking the official EC-6 TExES certification examination will not be the same for each student as the course BIOL 1082 can be taken at any time during the preservice teachers' enrollment in the EPP and memory decay varies from student to student.

Future Research and Suggestions

The findings of this study have valuable suggestions that may help EPPs to successfully identify preservice teachers who may need timely interventions before taking their official state certification examination. Previous studies explored possible factors that may influence students' performance on an examination however, there are limited studies that have been conducted on the performance of preservice teachers (Warren, 2017). In addition, the use of predictive modeling could be used with other mandatory science courses needed to complete the EPPs curriculum to make this study more generalized.

The mandatory science courses can be taken at any time during preservice teachers' tenure in the teacher preparation program (TPP). It is suggested that perhaps the order in which the courses are taken be rearranged to allow the courses that address the most competencies on the science portion of the official EC-6 exam, be taken closer towards the end of their program. A practice exam can be used to assess testing readiness.

The authors received no financial support for the research, authorship, and/or publication of this manuscript.

Ruthanne Thompson (Ruthanne.thompson@unt.edu) conducts research in three areas: pre-service science teacher education, science teacher retention, and undergraduate STEM retention. Dr. Thompson has 30+ publications including book chapters, textbooks, and research reports; is a University Distinguished Teaching Professor, and she and her lab have received two prestigious awards and over $15M in grant funding to support pre-service teacher education and STEM retention.

Naudin Alexis (naudinalexis@my.unt.edu) is a data analyst with Frisco ISD in Assessment and Accountability. She received a bachelor's degree in biology from Midwestern State University, a master's degree in biochemistry and molecular biology from the University of North Texas, and a Ph.D. in environmental science with focus in science education from the University of North Texas. Her research interests include scientific research in education and comparison of accountability standards between Education Preparation Programs (EPPs) and academic accountability for public school districts.

Pamela Esprivalo Harrell (pam.harrell@unt.edu) is a science educator whose research agenda demonstrates a continued focus on teacher quality. While at UNT, she worked with colleagues to launch one of the first three online teacher certification programs in the US that continues to enroll and graduate a high number of master's students. She received 32 grants totaling over $10M and served as Pl or CoPI for over $8M of grant funding. She created The Simply Outrageous Science Professional Development Program that served over 400 middle school science teachers, increasing their pedagogical knowledge and raising student achievement in high-need schools.

References

Alismail, H. A. (2016). Multicultural education: Teachers' perceptions and preparation. *Journal of Education and Practice, 7*(11), 139-146.

Acharya, B. R. (2017). Factors affecting difficulties in learning mathematics by mathematics learners international. *Journal of Elementary Education, 6*(2), 8-15.

Al Sheeb, B., Abdella, G.M., Hamouda, A.M., & Abdulwahed, M.S. (2019). Predictive modeling of first-year student performance in engineering education using sequential penalization-based regression. *Journal of Statistics and Management Systems, 22*(1), 31-50.

Bains, L. (2011). An Analysis OF TExES Sores: Evaluating Teacher Preparation at one Institution in the State of Texas. Unpublished doctoral dissertation, University of Houston, Houston, TX.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York, NY: W. H. Freeman and Company.

Banks, T., Jackson, D., & Harper, B. (2014). Responding to the call to prepare highly effective teacher candidates in the United States: The curriculum redesign effort in advancing teacher education. *Higher Education Studies*, *4*(2), 9-18.

Binns, I. C., Casto, A. R., Polly, D., & Bickmore, S. T. (2020). Examining pre-service teachers' integration of science and literacy through trade books. *Electronic Journal for Research in Science and Mathematics Education*, *24*(2), 7-21.

Bird, K.A., Castelman, B.L., Mabel, Z., & Song, Y. (2021). Bringing Transparence to Predictive Analytics: A systematic Comparison of Predictive Modeling Methods in Higher Education. *AERA Open, 7*.

Boyd, D., Goldhaber, D., Lankford, H., & Wyckoff, J. (2007). The effect of certification and preparation on teacher quality. *The Future of Children*, *17*(1), 45-68.

Capraro, R. M., Capraro, M. M., Parker, D., Kulm, G., & Raulerson, T. (2005) The Mathematics Content Knowledge Role in Developing Preservice Teachers' Pedagogical Content Knowledge. *Journal of Research in Childhood Education*, *20*(2), 102-118.

Chao-Ying, J. P., & Tak-Shing H. S. (2002). Logistic regression analysis and reporting: A primer. *Understanding Statistics, 1*(1), 31-70.

Collister, B. & Ellis, J. (2015). Central Texas teachers often struggle becoming certified. KXAN Investigates. http://kxan.com/investigative-story/central-texas-teachers-often-strugglebecoming-certified.

Cook, P. R. (2018). *An Exploration of Educational Trends (V2): A Symposium in Belize, Central America,* 110-115

Corcoran, R. P., & O'Flaherty, J. (2018). Factors that predict pre-service teachers' teaching performance. *Journal of Education for Teaching*, *44*(2), 175-193

D'Amico, M.M., & Dika, S.L. (2013). Using data known at the time of admission to predict first-generation college student success. *Journal of College Student Retention: Research, Theory & Practice, 15*(2), 173-192.

Darling-Hammond, L., Bae, S., Cook-Harvey, C.M., Lam, L., Mercer, C., Podolsky, A., & Stosich, E. L. (2016). Pathways to New Accountability Through Every Student Succeeds Act *Learning*

Policy Institute. Retrieved from https://learningpolicyinstitute.org/sites/default/files/product-files/Pathways_New-Accountability_Through_Every_Student_Succeeds_Act_04202016.pdf

Darling-Hammond, L. (2019). *A license to teach: Building a profession for 21st century schools*. Routledge.

DeFreitas, S. C. & Rinn, A. (2013). Academic achievement in first generation college students: The role of academic self-concept. *Journal of the Scholarship of Teaching and Learning, 13*(1), 57-67.

Donovan, C. B., Ashdown, J. E., & Mungai, A. M. (2014). A new approach to educator preparation evaluation: Evidence for continuous improvement? *Journal of Curriculum and Instruction*, *8*(1), 86-100.

Douglass, J., & Thomson, G. (2012). Poor and rich: Student economic stratification and academic performance in a public research university system. *Higher Education Quarterly, 66*(1), 65-89.

Every Student Succeeds Act, 20 U.S.C. § 6301 (2015). https://www.congress.gov/bill/114th-congress/senate-bill/1177

Feuer, M. J., Floden, R. E., Chudowsky, N., & Ahn, J. (2013). Evaluation of Teacher Preparation Programs: Purposes, Methods, and Policy Options. National Academy of Education. 500 Fifth Street NW Suite 339, Washington, DC 20001.

Frizzell, M., & Moffett, N.L. (2014). The Identification of Variables and Factors Related to Preservice Teacher Candidates' Passing a State Teacher Certification Examination at an HBCU.

Gard, D. R. (2011). *Teacher Certification Exams: Predicting Failure on the TExES History (8-12) Content Exam (A Nonparametric Approach using Classification Trees)*. Unpublished doctoral dissertation, Texas Tech University, Texas.

Graunke, S.S., & Woosley, S.A. (2005). An exploration of the factors that affect the academic success of college sophomores. *College Student Journal, 39*(2), 367-377.

Hosmer, D.W. & Lemeshow, S. (2000). *Applied logistic regression* (2nd edition). John Wiley & Sons, Inc., New York. doi:10.1002/0471722146

Huang, S., & Fang, N. (2013). Predicting student academic performance in an engineering dynamics course: A comparison of four types of predictive mathematical models. *Computers & Education, 61*, 133-145.

Hung, J.L., Shelton, B.E., Yang, J., & Du, X. (2019). Improving predictive modeling for at-risk student identification: A multistage approach. *IEEE Transactions on Learning Technologies, 12*(2), 148-157

Hutson, T. (2017). Preparing for EC-6 Teacher Certification in Texas: How much Science is Enough? In P. R. Cook (Ed.), *An Exploration of Educational Trends (V2): A Symposium in Belize. Central America* (pp.111-120). United Kingdom: Cambridge Scholars Publishing.

Hyndman, R. J. and Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*(4), 679-688.

Karamazova, E., Zenku, T. J., & Trifunov, Z. (2017). Analyzing and comparing the final grade in mathematics by linear regression using Excel and SPSS. *International Journal of Mathematics Trends and Technology (IJMTT), 52*(5), 334-344.

Kazempour, M., & Sadler, S. D. (2015). Pre-service teachers' science beliefs, attitudes, and self-efficacy: a multi-case study. *Teaching Education, 26*(3), 247-271.

Khajuria, S. (2007). *A model to predict student matriculation from admissions data*. Unpublished master's thesis, Russ College of Engineering and Technology, Ohio University, Ohio.

Kim, E., & Corcoran, R. P. (2018). Factors that influence pre-service teachers' persistence. *Teaching and Teacher Education*, *70*, 204-214.

Krieg, J.M. (2010). Academic performance of native and transfer students.

Kuha, J., & Mills, C. (2017). On group comparisons with logistic regression models. *Sociological Methods & Research*. ISSN 00491241

Lesley, E. L. (2016). What parents still do not know about No Child Left Behind and why it matters. *Journal of Education Policy*, *31*(3), 343-361.

Linn, V., & Jacobs, G. (2015). Inquiry-based field experiences: transforming early childhood teacher candidates' effectiveness. *Journal of Early Childhood Teacher Education*, *36*(4), 272-288.

Lykourentzou, I., I. Giannoukos, G. Mpardis, V. Nikolopoulos, & V. Loumos. (2009). Early and dynamic student achievement prediction in e-learning courses using neural networks. *Journal of the American Society for Information Science and Technology*, *60*(2), 372-380.

Martinez, J.A., Sher, K.J., Krull, J.L., & Wood, P.K. (2009). Blue-collar scholars?: Mediators and moderators of university attrition in first-generation college students. *Journal of college student development, 50*(1), 87.

Masters, G. (2018). The role of evidence in teaching and learning. *Australian Council for Educational Research*, *4*, 3-5.

Miller, S. P & Hudson, P. J. (2007). Using evidence-based practices to build mathematics competence related to conceptual, procedural, and declarative knowledge. *Learning Disabilities Research and Practice*, *22*(1), 47-57.

Murphy, C., Neil, P., & Beggs, J. (2007). Primary science teacher confidence revisited: ten years on. *Educational Research*, *49*(4), 415-430.

Murphy, C., & Smith. G. (2012). The impact of a curriculum course on pre-service primary teachers' science content knowledge and attitudes towards teaching science. *Irish Educational Studies*, *31*(1), 77-95.

NCES. (2017). The Condition of Education. Washington, D.C.: National Center for Educational Statistics.

NCTQ. (2014). Teacher Prep Review. Washington, D.C.: National Council on Teacher Quality.

Peters, B. A., & Draughon, G. B. (2017). Part-Time Students and Performance-Based Funding 2.0 in United States Public Higher Education. In H. Alphin, Jr., J. Lavine, and R. Chan (Ed.), Disability and Equity in Higher Education Accessibility (pp. 153-176). IGI Global. http://doi:10.4018/978-1-5225-2665-0.ch007

Pino-Pasternak, D., & Volet, S. (2018). Evolution of pre-service teachers' attitudes towards learning science during an introductory science unit. *International Journal of Science Education*, *40*(12), 1520-1541.

Sandholtz, J. H., & Shea, L. (2015). Examining the Extremes: High and Low Performance on a Teaching Performance Assessment for Licensure. *Teacher Education Quarterly,* 42(2), 17–42.

Sasson, I. (2014). The role of informal science centers in science education: Attitudes, skills, and self-efficacy. *Journal of Technology and Science Education*, *4*(*3*), 167-180.

Senler, B. (2016). Pre-service science teachers' self-efficacy: The role of attitude, anxiety, and locus of control. *Australian Journal of Education*, *60*(1), 26-41.

Shipe, M. E., Deppen, S. A., Farjah, F., & Grogan, E. L. (2019). Developing prediction models for clinical use using logistic regression: an overview. *Journal of thoracic disease*, *11*(4), 574–584. https://doi.org/10.21037/jtd.2019.01.25

Smith, T.J., & McKenna, C.M. (2013). A comparison of logistic regression pseudo R2 indices. *Multiple Linear Regression Viewpoints, 39*(2), 17-26.

Smith, V.C., Lange, A., & Huston, D.R. (2012). Predictive modeling to forecast student outcomes and drive effective interventions in online community college courses. *Journal of asynchronous learning networks, 16*(3), 51-61.

Sperandei, S. (2014). Understanding logistic regression analysis. *Biochemia Medica*, *24*(1), 12-18.

Stroud, J. C., Smith, L. L., Ealy, L. L., & Hurst, R. (2000). Choosing to Teach: Perceptions of Male Preservice Teachers in Early Childhood and Elementary Education. *Early Child Development and Care*, *163*(1), 49-60.

Sullivan, W. G., Daghestani, S. F., & H. R. Parsaei. (1996). Multivariate analysis of student performance in large engineering economy class. *Technology-Based Re-Engineering Engineering Education Proceedings of Frontiers in Education FIE'96 26th Annual Conference*, *1*, 180-184.

Sutcher, L., Darling-Hammond, L., & Carver-Thomas, D. (2016). *A coming crisis in teaching? Teacher supply, demand, and shortages in the U.S.* Palo Alto, CA: Learning Policy Institute., 1-16. Retrieved May 17, 2018, from https://learningpolicyinstitute.org/sites/default/files/product-files/A_Coming_Crisis_in_Teaching_BRIEF.pdf

Swecker, H. K., Matthew, F., & Searby, L. (2013). Academic advising and first-generation college students: A quantitative study on student retention. *NACADA Journal*, *33*(1), 46- 53.

Texas Education Agency (TEA). (2018). *Educator certification.* Retrieved from http://tea.texas.gov/Texas_Educators/

Thobega, M, & Miller, G. (2008). Predicting Scores of Beginning Agricultural Education Teachers on the Praxis II Examination. *Journal of Agricultural Education*, *49* (1), 99–109.

Van Aalderen-Smeets, S. I., Van der Molen., Van Hest, J. H. W., & Poortman, C. (2017). Primary teachers conducting inquiry projects: effects on attitudes towards teaching science and conducting inquiry. *International Journal of Science Education*, *39*(2), 238-256.

Warren, S. (2017). *The relationship between admission criteria and TExES Generalist EC-6: An archival study of predictors of future success in the teacher preparation program at a historical black college in Southwest Texas.* Unpublished doctoral dissertation, North Central University, Minneapolis, MN.

Yang, S. J. H., Lu, O. H. T., Huang, A. Y. Q., Huang, J. C. H., Ogata, H., & Lin, A. J. Q. (2018). Predicting students' academic performance using multiple linear regression and principal component analysis. *Journal of Information Processing*, *26*, 170-176.

Appendix A

Biology for Educators Survey

Q1 Last name

____

Q2 First name

____

Q3 Student ID number

____

Q4 UNT Email address

____

Q5 Gender

- ▢ Male
- ▢ Female
- ▢ Do not wish to share

Q6 Age

- ▢ 18-20
- ▢ 21-23
- ▢ 24-26
- ▢ 27-30
- ▢ 31-35
- ▢ 36-40
- ▢ 40+

Q7 What is your current classification?

- ▢ Freshman
- ▢ Sophomore
- ▢ Junior
- ▢ Senior
- ▢ Post Baccalaureate
- ▢ Graduate student
- ▢ Other ____

Q8 From the following choices, please select the type of high school you attended. In the box under the type of high school, please write the name of the high school, and the school district if applicable.

- ☐ Public school ______________________________
- ☐ Private school ______________________________
- ☐ Boarding school ______________________________
- ☐ Home schooled

Q9 Check the box next to all the science courses you have taken in high school. If some of the science course/s were taken at a different high school/s, write name of school next to the name of the science course. If applicable, indicate if the course was an honor/advanced placement course in the box next to the science course. If the course is not listed, please check the box "other" and write the name/s of the course/s in the box.

- ☐ Biology ______________________________
- ☐ Chemistry ______________________________
- ☐ Physics ______________________________
- ☐ Environmental science______________________________
- ☐ Aquatic science ______________________________
- ☐ Forensic science ______________________________
- ☐ Earth science ______________________________
- ☐ Agriculture ______________________________
- ☐ Marine biology ______________________________
- ☐ Botany ______________________________
- ☐ Zoology ______________________________
- ☐ Other ______________________________

Q10 Are you a transfer student? If so, please state the college and/or university you transferred from. continue to the following question to select the science courses you have taken at the college level prior to attending UNT.

- ☐ Yes______________________________
- ☐ No

Q11 If you have transferred from a college and/or university, please select from the list below the science courses you have taken there. If applicable, please write the name of college or university where the course was taken in the box next to the course

- ☐ Biology ______________________________

- ☐ Chemistry ______________________________
- ☐ Physics ______________________________
- ☐ Environmental science ______________________________
- ☐ Aquatic Science ______________________________
- ☐ Forensic science______________________________
- ☐ Earth science ______________________________
- ☐ Agriculture ______________________________
- ☐ Marine biology ______________________________
- ☐ Botany ______________________________
- ☐ Zoology ______________________________
- ☐ Other ______________________________

Q12 Check the box next to all science courses that you have taken at UNT prior to taking this class. State the name of college where the science course/s was taken next to the science course. If applicable, indicate if the course was an honor/advanced placement course in the box next to the science course. If the course is not listed, please check the box "other" and write the name/s of the course/s in the box.

- ☐ Biology ______________________________
- ☐ Chemistry ______________________________
- ☐ Physics ______________________________
- ☐ Environmental science ______________________________
- ☐ Aquatic Science ______________________________
- ☐ Forensic science______________________________
- ☐ Earth science ______________________________
- ☐ Agriculture ______________________________
- ☐ Marine biology ______________________________
- ☐ Botany ______________________________
- ☐ Zoology ______________________________
- ☐ Other ______________________________

Q13 List the science courses that you are currently enrolled in while taking this class. If the course is not taken at UNT, please write the name of the college next to the course. If applicable, indicate if the course is an honors course.

- ☐ Course 1 ______________________________

☐ Course 2 ____

☐ Course 3 ____

☐ Course 4 ____

☐ Course 5 ____

Q14 Are you a part-time or a full-time student? List the number of credits you are currently taking next to your choice.

☐ Part time ____

☐ Full time ____

Q15 Are you a first-generation college student? Meaning are you the first member of your family to attend college?

Optional: If no, please state your relationship to the person who previously attended.

☐ Yes

☐ No ____

Q16 Do you have a parent and/or a close relative who obtained a degree in Education?

☐ Yes ____

☐ No

Q17 Why did you decide to pursue an Education degree?

____

Q18 What grade level do you plan on teaching?

____

Q19 Have you taken this class before? If yes, what grade did you receive and what is the reason for retaking this class?

o Yes ____

o No

20 What is your current GPA?

☐This is permission to access your GPA from the Office of Institutional Research at UNT. Write your first name and last name in box below for full consent.

____

☐ GPA can be accessed at my.unt.edu

____

Q21 Check the box of the standardized tests listed below that you have taken. Write the score you obtained (to the best of your recollection) in the space under the name of the test.

☐This is permission to access any of the below scores that are available from the Office of Institutional Research at UNT. Write your first name and last name in box below for full consent.

____

- ▢ ACT or SAT ____
- ▢ GRE ____
- ▢ TAKS or STAAR (science portion) or if high school is out Texas, the equivalent end of course exit exam ____
- ▢ Other? Write the name of the test and score (optional) below.

____