Technology as a Support for Proof and Argumentation: A Systematic Literature Review

By Tye G. Campbell[1] and Jeremy Zelkowski[2]

[1] University of Alabama, Tuscaloosa, AL, USA; ORCID: 0000-0001-7827-8430
[2] University of Alabama, Tuscaloosa, AL, USA

tgcampbell1@crimson.ua.edu, jzelkowski@ua.edu

Received: 14 March 2019 Revised: 1 July 2019 **DOI: 10.1564/tme_v27.2.04**

Proof and argumentation are essential components of learning mathematics, and technology can mediate students' abilities to learn. This systematic literature review synthesizes empirical literature which examines technology as a support for proof and argumentation across all content domains. The themes of this review are revealed through analyzing articles related to Geometry and mathematical content domains different from Geometry. Within the Geometry literature, five subthemes are discussed: (1) empirical and theoretical interplay in dynamic geometry environments (DGEs), (2) justifying constructions using DGEs, (3) comparing technological and non-technological environments, (4) student processing in a DGE, and (5) intelligent tutor systems. Within the articles related to content different from Geometry, two subthemes are discussed: technological supports for number systems/algebra and technological supports for calculus/real analysis. The technological supports for proof revealed in this review could aid future research and practice in developing new strategies to mediate students' understandings of proof.

Keywords: Proof and proving, Argumentation, Technology, Dynamic Geometry Software

1. INTRODUCTION

Proof and argumentation are important process standards for learning, doing, and understanding mathematics (Knuth, 2002). In fact, some scholars claim that proof lies at the core of mathematics and cannot be separated from the subject itself (Schoenfeld, 1994). Despite the importance of proof in learning mathematics, research consistently shows that students of all ages struggle to construct viable mathematical arguments (e.g. Healy & Hoyles, 2000; Lannin, 2005; Lin, Yang, & Chen, 2004; Sen & Guler, 2015). Because of this, extensive research is devoted to analyzing learners' struggles with proof and enhancing their argumentative capacities (e.g. Styliandies, Bieda, & Morselli, 2016; Stylianides, Stylianides, & Weber, 2017). Technology is of specific interest as a tool for mediating learners' capacities with proof and argumentation because there is a general consensus that technology enhances mathematical learning (National Council of Teachers of Mathematics [NCTM], 2000; UK Department of Education, 2014). Current literature reviews provide important contributions by synthesizing research related to Dynamic Geometry Software as a tool for proving within Geometry (Hollebrands, Laborde, & Sträßer, 2008; Sinclair & Robutti, 2013; Sinclair, et al., 2016). However, current reviews have not utilized a systematic methodology for searching for and finding literature and have only reported on one subject domain (Geometry). A systematic methodology ensures a broad coverage of the literature that might have been missed in previous reviews. Additionally, given an international emphasis on reasoning and proof in school mathematics across all content domains (e.g. NCTM, 2000; UK Department of Education, 2014), the field could benefit from a review of technological supports related to all mathematical subjects. To add to the current body of knowledge, we utilize a systematic methodology to find articles which report on technology as a support for engaging with proof and argumentation across all content domains.

In what follows, we detail our deliberate and systematic approach for including or excluding articles in this synthesis of the literature. However, we must first admit a partiality in our conceptualization of proof and argumentation. Harel and Sowder (1998) viewed proof as a process of ascertaining and persuading. In other words, proof and argumentation involves arguing for or against a mathematical claim and convincing oneself and others of the truth or falsity of the claim. We align with this view of proof and argumentation, though we acknowledge the differences in how others in the field conceptualize these terms. Additionally, some scholars acknowledge a primary distinction between the constructs of proof and argumentation. While we acknowledge potential differences in interpreting proof and argumentation, we align with Stylianides, et al.'s (2016) contention that "(1) argumentation and proof are closely related, and (2) considering both argumentation and proof helps draw attention to a wider range of important processes related to proving than when considering them separately" (p. 316). Therefore, both argumentation and proof are considered in this review. We made every effort to account for differing perspectives of proof and argumentation by including articles from a variety of conceptualizations. However, due to our biased interpretation of proof and argumentation, it is possible that we did not account for some publications which make an important contribution to the field.

2. METHODS

A systematic review methodology (Cooper, 2017; Hannes, Claes, & Belgian Campbell Group, 2007) was utilized to find and analyze research studies related to technology as a support for proof and argumentation. A systematic literature review follows a strategic process wherein articles are retrieved from databases using a bank of search terms, selected or excluded based on pre-determined criteria, and synthesized to portray themes of the literature. To determine which databases to use in our search, we consulted a research librarian who recommended three databases relevant to our research topic: ERIC EBSCOhost, Education Full Text (H.W. Wilson), and PsychINFO. Furthermore, we chose to review articles published after the year 2000 because NCTM's (2000) curriculum document, having an international impact, prioritized proof as a process standard for K-12 mathematics. Using limiters to include peer reviewed articles published after January 1, 2000, we conducted a simultaneous search in each of the three databases on January 3, 2019 including the following terms anywhere in the article: (proof OR argumentation) AND (math*[1]) AND (tech*). The terms were chosen based on curriculum documents in the field of mathematics education related to proof and technology (e.g. NCTM, 2000; UK Department of Education, 2014).

The simultaneous search yielded 938 articles in total with 754 articles remaining after duplicates were eliminated. Each of the 754 articles were transported to RefWorks, a tool for creating bibliographies, and they were subsequently transported to a spreadsheet for screening purposes. The criteria for inclusion in the review were as follows: (1) *publication type*—the publication was a journal article or conference paper, (2) *empirical*—the study reported empirical findings, (3) *participants*—the study included learners[2] of any age as participants, (4) *content*—the study focused on technology as a support for proof and/or argumentation, and (5) *language*—the study was written in English.

In the first phase of the screening process, the title of each article was examined to determine whether it met inclusion criteria. After the title screen, 207 articles remained. Next, each abstract of the remaining articles was reviewed to determine whether it met the inclusion criteria eliminating another 142 articles. Finally, the full text of each remaining article was scanned, and after this last phase of screening, 28 articles remained for inclusion. To seek possible publications that were missed in the search, an ancestral search was run on each of the references for the 28 included articles. The ancestral search yielded an additional 4 articles leaving a total of 32 articles included in this review. A PRISMA diagram (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group) of our search and selection process can be found in Figure 1.

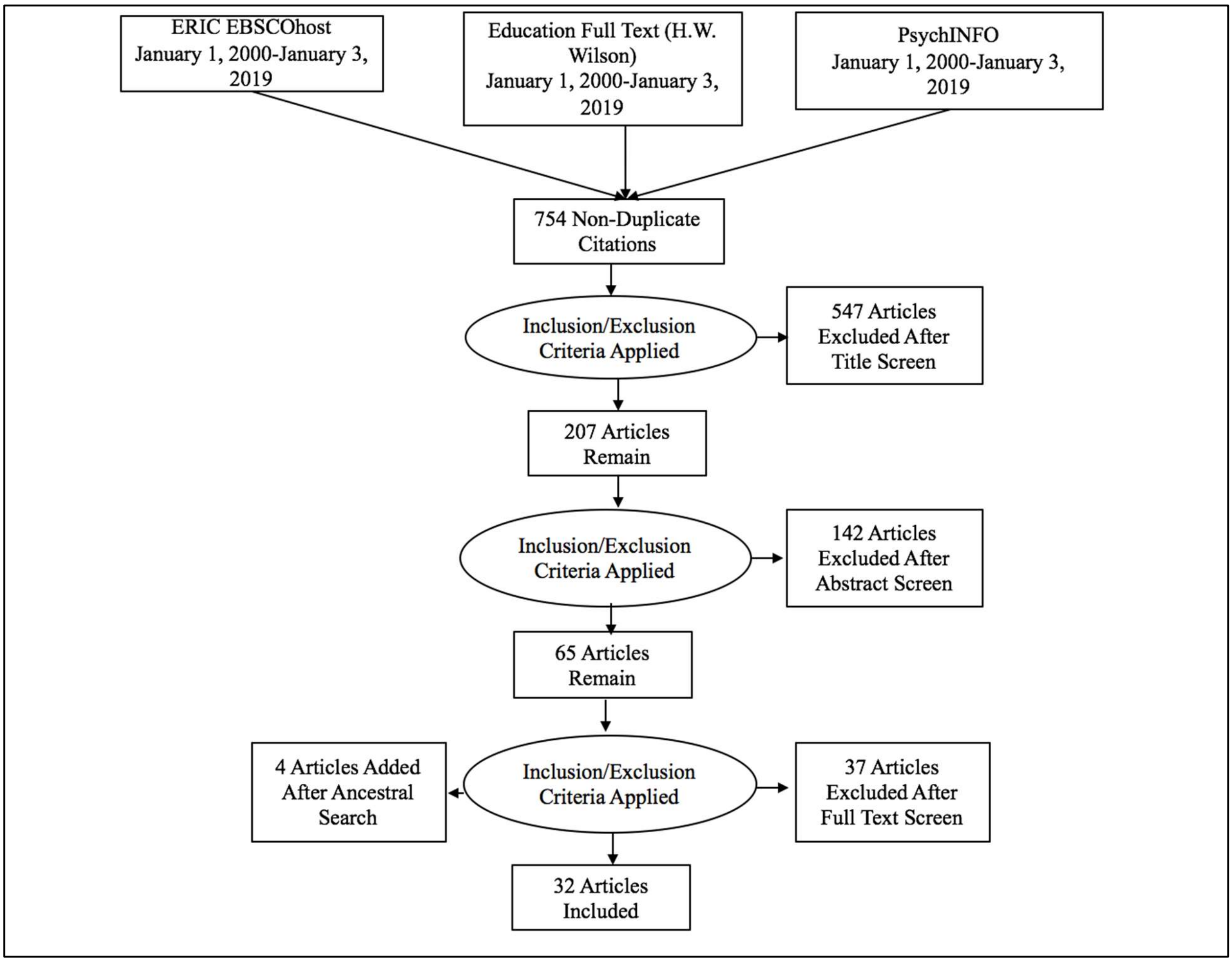

Figure 1 PRISMA Diagram

[1] Any letters occurring after * were included in the search.

[2] We did not consider studies including teachers as participants unless the teachers were situated as learners within a college course.

3. ANALYSIS

To analyze and synthesize the literature, each article was read in full and coded according to a rigorous coding scheme. Articles were coded for the number of participants, content of the mathematical task(s), theoretical framing, technological support(s), findings, among many other indicators which can be found in Table 1. After the coding process, the articles were analyzed according to a constant comparative method (Glaser & Strauss, 1967) to organize the literature into categories for discovering themes and sub-themes. The categories and themes constantly evolved during the analysis until an amenable set of themes were developed.

Our analysis revealed 22 of the 32 articles included in this review explored technology as support for proof in Geometry. Therefore, the first theme constitutes technological supports for proof and argumentation in Geometry while the second theme represents technological supports for proof and argumentation in subjects different from Geometry. We developed sub-themes within each of these major themes, and they are discussed further in the findings. Prior to sharing the thematic findings, we briefly share contextual features of the studies reported in this review based on the country in which the study took place, the number of participants, and the duration of the study. This information might provide the reader with an overall sense of the context of each article, and it might suggest implications for future research. In what follows, we present the findings followed by a discussion on implications for future research and practice.

What country did the study take place?
What was the participants' grade level? 1=K-2, 2=3-5, 3=6-8, 4=9-12, 5=College
What was the sample size? 1=30 or less, 2=30-60, 3=60-90, 4=90 or more
What content was used to solicit argumentation? 1=Algebra, 2=Number Theory, 3=Geometry, 4=Precal/Trig/Cal 5=Other (specify), 6=Unknown
What was the duration of the study? 1=One day or less, 2=1 week or less, 3=one month or less, 4=one marking period or less, 5=one semester (or summer) or less, 6=one school year or less, 7=more than one school year (specify)
Explain theoretical/conceptual framework used
Specify the technological support to solicit argumentation
How were students working on argumentative tasks? 1=individually, 2=collaboratively, 3=NA
Specify the methods used
Explain the findings

Table 1 Coding Scheme

4. FINDINGS

4.1 Contextual Features of Literature

Tables 2-4 present the statistical findings of the contextual features of the studies reported within this paper as a whole. Table 2 reveals more studies were conducted in the United States (8 out of 32) than any other country, though only including articles written in English could have impacted these results. European (Italy, Spain, Germany, United Kingdom) and Middle Eastern (Israel, Turkey, and Cyprus) countries also contributed much of the literature. Table 3 reveals the majority of the studies included less than 30 participants (21 out of 32), although almost 20% of the studies included 91 or more participants. Table 4 reveals the majority of studies were shorter than one week (18 out of 32), while some studies were more longitudinal in nature with seven studies extending more than one semester.

Country	Percentage of Publications
United States	25% (8/32)
Italy	16% (5/32)
Spain	13% (4/32)
Germany	9% (3/32)
Israel	9% (3/32)
Turkey	9% (3/32)
United Kingdom	9% (3/32)
Hong Kong	6% (2/32)
Cyprus	3% (1/32)
Taiwan	3% (1/32)

Note: Percentages exceed 100% because some studies took place in more than one country.

Table 2 Publications by Country

Number of Participants	Percentage of Publications
30 or less	66% (21/32)
31 to 60	13% (4/32)
61 to 90	3% (1/32)
91 or more	19% (6/32)

Table 3 Number of Participants

Duration of the Study	Percentage of Publications
One day or less	22% (7/32)
More than one day and less than one week	34% (11/32)
More than one week and less than one month	22% (7/32)
More than one month and less than one semester	16% (5/32)
More than one semester and less than one year	3% (1/32)
More than one year	3% (1/32)

Table 4 Duration of the Study

4.2 Technological Supports in Geometry

Through synthesizing the literature, five sub-themes related to technology as a support for proof and argumentation in Geometry were constructed: (1) empirical and theoretical interplay in dynamic geometry environments (DGEs), (2) justifying constructions using DGEs, (3) comparing technological and non-technological environments, (4) student processing in a DGE, and (5) intelligent tutor systems. All of the articles used DGEs as the technological support for supporting proof and argumentation signifying the importance scholars place on this specific tool. Mariotti (2001) explained the inner-workings of one DGE called Cabri-Géomètre as a "microworld which embodies Euclidean geometry, with its elements and its properties (points, line, circles, but also midpoint, angle bisector, perpendicularity, parallelism, . . .)" (p. 260). Though there are many types of DGEs, they are all equipped with several tools that help students construct and manipulate figures, check measurements, explore geometric properties on a computerized screen, and test multiple cases quickly and efficiently. Mariotti (2000) contends "the novelty of a dynamic environment consists in the direct manipulation of its figures" (p. 27). That is, all manipulations of a figure occur in real time allowing the student to visualize geometric properties dynamically rather than from a static orientation. The thematic findings that follow further provide a comprehensive description of DGEs.

4.2.1 Empirical and Theoretical Interplay in DGEs

The literature strongly suggests that technology supports students' argumentative capacities in Geometry when the tools are used effectively. However, several scholars caution that technology, and particularly DGEs, might lead students to become less appreciative of the necessity of proof (e.g. Christou, Mousoulides, Pittalis, & Pitt-Pantazi, 2004; Lachmy & Koichu, 2014; Mariotti, 2002; Marrades & Gutiérrez). This is because in a DGE, students can empirically test mathematical claims using dragging functions or measuring tools and become convinced of the veracity of a claim without deductive justification (Lachmy & Koichu, 2014). Leung & Lopez-Real (2002) capture the tension between the deductive nature of geometry and the empirical qualities of DGE in their assertion "when this empirical and inductive dimension is to be added to a pedagogical structure that is traditionally rooted in deductive logic, careful examination is needed on how to combine these two seemingly opposite perspectives" (p. 4). Because of this tension, several scholars explored how to support students in connecting empirical investigations in DGE with theoretical geometric proof.

Some scholars suggest the design of a task could help students see the connection between empirical investigations and theoretical proof (Christou, et al, 2004; Guven & Karatas, 2009; Hadas, Hershkowitz, & Shwarz, 2000). Christou, et al. (2004) designed a task for three pre-service primary school teachers specifically asking them to explore within Geometer's Sketchpad, create a conjecture, and explain why the conjecture is true. Within their task sequence, they found pre-service teachers felt more compelled to prove the conjecture because they had the opportunity to discover within the DGE and create their own conjecture instead of proving a given claim. Similarly, Hadas, et al. (2000) designed an open task sequence to help students move from empirical aspects of DGE to theoretical grounds. In their study with eighth and tenth grade participants, they created a task wherein students created a conjecture on the sums of interior and exterior angles of a polygon and explained why their conjecture was true. However, Hadas, et al. (2000) specifically designed the task to lead students to a false conjecture. After being surprised that their conjecture was incorrect, students understood that proof was necessary even after using tools within the DGE to test their conjectures. Additionally, 50% of the eighth-grade students and 56% of tenth-grade students had success in creating a viable argument after changing their conjecture. Guven & Karatas' (2009) found pre-service teachers engaged in a mathematization process of experimenting, conjecturing, and proving within a DGE when they were given open tasks that allowed for multiple conjectures and points of entry. Other scholars (e.g. Healy & Hoyles, 2002; Marrades & Gutiérrez, 2000) similarly noted the importance of the conjecture process before engaging in proof.

Instead of focusing on the task sequence, Leung and Lopez-Real (2002) sought to understand how two students, Hilda and Jane, operated between empirical and theoretical grounds when doing a proof by contradiction using Cabri Geometry Software. Through their observations of Hilda and Jane, Leung and Lopez-Real (2002) developed a framework for the argumentative stages of proof by contradiction in DGE. First, the student creates a biased DGE microworld followed by constructing a pseudo-object, or an impossible Euclidean figure. Then, the student discovers a locus of validity, or a confinement through which the pseudo-object is valid. Lastly, the student makes a conjecture and organizes their proof by contradiction. Lueng and Lopez-Real (2002) suggested the critical stage of creating a pseudo-object "might bring about the cognitive unity of a theorem bridging the empirical-theoretical gap between inductive acquisition and formal justification" (p. 21). Building on Lueng & Lopez-Real's (2002) work, Baccaglini-Frank, Antonini, Leung, & Mariotti (2013) developed the notion of a proto-pseudo object which is "a geometrical object that has the potential of becoming a pseudo object" (p. 65). The transition from a proto-pseudo object to a pseudo object proved essential to having success in proof by contradiction in their analysis. Together, these studies revealed that an appropriate task sequence and careful attention to student reasoning can overcome the divide between empirical aspects of DGE and theoretical features of Euclidean Geometry.

4.2.2 Justifying Constructions using DGEs

Other scholars focused on constructions within a DGE and how students justified their constructions. Mariotti's (2000) study with high school students showed they were able to create theoretical rather than intuitive understandings of constructing geometric figures as a result of using Cabri Geometry Software. At the beginning of their investigations, students saw dragging as an external process to constructing a figure. However, through mathematical discussions and social learning with peers and the instructor, students began to see dragging as a theoretical move in the construction of the figure. In other words, the geometry software became part of the theory for proving rather than a simple tool. In a later study with ninth and tenth-grade students, Mariotti (2002) developed a sequence of instruction for using Cabri Geometry Software to aid students in justifying their conjectures. First, she suggested that teachers should encourage students to explain their reasoning throughout the construction process. Then, through a negotiation process, students should make their constructions acceptable to the classroom community. Lastly, students should adapt their justifications to defend their constructions.

Similarly, Jones (2000) engaged 26 twelve year-old students in a three-phase sequence wherein students justified their geometric constructions within Cabri-Géomètre. In the first phase of instruction, students were to construct figures which were invariant under dragging. During this phase, students gave descriptions rather than justifying their conjectures. In the second phase, students were required to construct quadrilaterals that were invariant under dragging and explain why their construction produced the quadrilateral. In this phase, students' explanations were tied to the DGE and had few theoretical underpinnings. Instead, they used terms such as "dragging" as warrants for their claims. In the final phase, students were to create explanations about relationships between various quadrilaterals. By the third phase, the students created mathematical explanations that did not rely on the DGE. Jones' (2000) study showed that students require time to create deductive justification for their constructions. However, his study is an existence proof that students' mathematical arguments for constructions can be improved within a DGE. The literature on justifying constructions suggest DGEs can be a powerful way for helping students create theoretical understandings for their constructions.

4.2.3 Comparing Technological and Non-technological environments

Several studies examined the differences between technological and non-technological settings on students' affect towards mathematics (Gómez-Chacón, Albaladejo, & López, 2014; Zengin, 2017) and their abilities to create deductive arguments (Hollebrands, Connor, & Smith, 2010; Smith, 2011). Gomez, et al. (2014) suggested that affect influences students' cognition as it relates to geometry proof. In their study with two classes of students aged 14-15 years old, participants engaged in a task sequence wherein the teacher taught 13 reasoning tasks using paper-and-pencil as tools and 12 reasoning tasks using GeoGebra. By the end of their task sequence, 42% of students created viable arguments using GeoGebra compared to only 9% using paper-and-pencil. Additionally, students showed much more perseverance when using GeoGebra as compared to working with paper-and-pencil. Gomez, et al. (2014) hypothesized that affect and cognition worked conjointly in students' reasoning processes while using GeoGebra. That is, students had more tools to persevere using a DGE and were more likely to try new strategies independently which caused them to have greater success in doing proofs. Zengin (2017) worked with pre-service mathematics teachers for 9 weeks working through reasoning and proof tasks in GeoGebra. Using an established Likert-scale measurement tool, he evaluated how students' attitudes changed towards proof as a result of using GeoGebra using a pre/posttest design. Higher scores on the test indicated positive attitudes while lower scores indicated negative attitudes. There was a significant difference in students' attitudes towards proof on the post-test (M=85.77) when compared with the pre-test (M=73.13), $t(21)=6.06$, $p<.05$, $r=.79$. This indicates the intervention had a large effect on students' attitudes towards proof. While it could be argued that their attitudes changed for reasons unrelated to GeoGebra, qualitative findings revealed students specifically acknowledged the DGE as a determinant for changing their views towards proof.

Two studies (Hollebrands, et al, 2010; Smith, 2011) found technology to enhance students' proof explorations when experimentally comparing the effects of DGE with non-technology environments. Hollebrands, et al. (2010) studied how eight college students used NonEuclid, a DGE for non-Euclidean Geometry, in their attempts to create proofs. They found students often used the software to check the validity of claims, and it was very beneficial in their investigations. However, students did not use explicit warrants for their arguments when working in the DGE. The authors hypothesized that students were accustomed to working on proofs with static figures rather than dynamic ones which might have made them unsure about their arguments within the DGE. In his study with eighth-grade students, Smith (2011) found students created more arguments when using Geometer's Sketchpad compared to an environment with physical manipulatives such as snap cubes. Students also frequently referred to the technology as warrants for their arguments in the DGE though they rarely referred to physical manipulatives as warrants in the non-technology environment. However, similar to Hollebrands, et al. (2010), Smith (2011) found students rarely used explicit warrants when working in the DGE. Hollebrands, et al.'s (2010) and Smith's (2011) studies revealed DGEs enhanced students' exploration processes but did not necessarily impact their abilities to create deductive arguments. Together, the research suggests DGEs have a positive effect on students' motivation and exploration of geometric proof when compared with non-technological environments.

4.2.4 Student Processing in a DGE

A few studies focused on students' processes while they worked within a technological environment on geometry proof tasks. Within a DGE, Lachmy and Koichu (2014) proposed the notion of direct variants (i.e. elements of a figure that may be directly manipulated through dragging) and indirect variants (i.e. elements of a figure that depend on dragging direct variants). Creating an argument in a DGE requires one to develop a hierarchy of dependencies relating direct and indirect variants. In Lachmy and Koichu's (2014) case study of one high attaining ninth grade student, they found the student had more success in creating arguments for conditional statements when compared with biconditional statements. The authors suggested that the student struggled constructing a figure in the DGE because she did not realize the appropriate hierarchy of dependencies for biconditional statements. Focusing on students' processes in an unmoderated DGE, Fukawa-Connelly and Silverman (2015) found that justifying a mathematical argument became normalized over time. In their study, three students used a software called Virtual Math Teams with GeoGebra for four weeks to construct figures, conjecture, and justify within a DGE equipped with a chat box available for students to converse with one another virtually. At the beginning of the intervention, students often solved tasks without feeling the need to justify their responses to their classmates . However, argumentation became a normalized practice by the end of the four weeks. The authors suggested the interaction between students was unmoderated, so the article provided little information for how the practice of argumentation became normalized over time. Soldano and Arzarello (2016) used game theoretic logic to design a game-type DGE called Geometric Constructor wherein students worked in dyads to attempt to create favorable situations to win the game. One student always had a more favorable position, and the students were to justify why one position was more favorable than the other. The authors found students processed reasoning and proof in three ways when using Geometric Constructor. Students used the reflected game to heuristically answer questions, provide evidence for a claim, and to check extreme cases. Each of these processes are important for developing deductive arguments, but the authors noted many students persisted in relying on empirical evidence. They concluded more research is needed to determine how and if games can help students create deductive arguments. While the current literature provides some insight into how students process within a DGE, there is much left to be discovered in this theme.

4.2.5 Intelligent Tutor Systems

Four studies explored DGEs equipped with intelligent tutor systems as a support for proof and argumentation. Paneque, Cobo, and Fortuny (2017) used GeoGebraTUTOR, a DGE that allows teachers to pre-load a software with possible solution paths that students might use when creating a proof, to determine the teacher-tutor interactions that were necessary to aid students in their deductive arguments. In their study, the teacher pre-loaded scaffolds into the DGE so a virtual tutor could respond with messages to re-direct students' thinking when they started an incorrect solution path. Through their observations of the teacher, tutor, and students, they noticed GeoGebraTUTOR helped students think strategically about proof if the scaffolds were presented in a way that maintained the challenge of the problem. That is, the teacher had to make sure their pre-loaded scaffolds appropriately guided the students without reducing the cognitive demand of the problem. Cobo, Fortuny, Puertas, and Richard (2007) used a similar DGE equipped with a virtual tutor called AgentGeom. AgentGeom is equipped with a graphic area where students can build geometric figures and a deduction editor where students can test hypotheses, check measurements, and validate other properties. The student may interact with the virtual tutor at any time by asking for help, or the tutor will provide help if the student starts down an incorrect solution path. In their case study of one 16 year-old named Gerard, they found Gerard appropriated the knowledge needed to create a viable argument through discursive actions with the artificial tutor. Both the graphic area and deductive area proved to be essential in Gerard's success at creating a viable argument.

Matsuda and VanLehn (2005) also used a DGE with a virtual tutor that used pre-loaded messages to interact with students. However, their study experimentally tested the effects of two scaffolds provided by the virtual tutor: forward chaining and backward chaining. Forward chaining refers to starting with a given statement and providing deductive warrants to arrive at a conclusion while backward chaining "starts from a goal to be proved and applies postulates backwards, that is, by matching a conclusion of the postulate to the goal, then posting the premises that are not yet proved as new goals to be proved" (p. 443). starting with the claim and working backwards deductively towards the given statement. In a pre/post-test experimental design with 52 undergraduate students, Matsuda and VanLehn (2005) found forward chaining was a more productive scaffold from the virtual tutor than backward chaining. However, both scaffolds improved students' argumentative abilities (Regression equation: Post-test=0.52*pre-test−0.14(if BC)+0.50). Wong, Yin, Yang, and Cheng (2011) similarly used an experimental design to determine the effects of an intelligent system called MR Geo (Multiple Representations for Geometry) on ninth-grade students' proving abilities. Instead of using a virtual tutor, MR Geo scaffolds instruction by providing the student with four representational systems on one screen. The four panes the student could interact with were the problem pane, the dynamic pane, the formal proof pane, and the proof tree pane. The student could work within each pane at the same time, and certain features in each pane would highlight to show the corresponding moves in another pane. Wong, et al. (2011) found that medium and low achievement groups (as determined by a pre-test and math grades) benefitted more than high achievement groups from their interaction with MR Geo. The four studies using intelligent tutor systems show positive outcomes for improving all students' proving capabilities, but lower-performing students might benefit more from these systems.

4.3 Technological Supports in Subjects other than Geometry

There were ten studies which used technology as a support for proof and argumentation in subjects other than Geometry. The ten articles were all related to number systems/algebra or calculus/real analysis which constitute the sub-themes for this section. In what follows, the findings are synthesized to report the current state of research in this domain.

4.3.1 Number Systems/Algebra

Three studies explored the same data set to determine the effects of collaboration scripts and heuristic worked examples on students' general and mathematical argumentative capacities while solving Number Theory tasks (Kollar, et al., 2014; Vogel, et al., 2016; Schwaighofer, et al., 2017). Collaboration scripts are virtual scripts that help students organize their group communication to solve a problem while heuristic worked examples are solved examples that show the strategies an imaginary student uses to solve a problem (Schwaighofer, et al., 2017). Heuristic worked examples are especially relevant to proof because proofs require students to learn strategies rather than procedures. They show the processes that a student might use to solve a specific proof problem, so students can extrapolate their understanding to a new task. In all three studies, the collaboration scripts and heuristic worked examples were computerized allowing the students to work between both scaffolds. In each of the three studies, 101 pre-service mathematics teachers were randomly placed in one of four conditions based on a 2x2 factorial (heuristic worked examples vs. no example support, collaboration script vs. no collaboration support), and they worked in dyads to complete argumentation tasks. Using a pre/post-test design, Vogel, et al. (2016) found the two supports had a positive effect on students' disposition to use general argumentation skills. Controlling for pre-test scores, Kollar, et al. (2014) similarly found groups using collaboration scripts made significantly higher gains in their social discursive argumentation skills (i.e. the ability to socially participate in argumentative practices) ($F(1,96)=4.42$, $p=.04$, part. $\eta^2=.04$), and there was also a positive effect for students who received heuristic worked examples compared to those who received no example support ($F(1,96)=9.68$, $p<.01$, part. $\eta^2=.09$). However, when specifically focusing on the mathematical component of argumentation, Kollar, et al. (2014) found students with low prior achievement performed better on the mathematical argumentation post-test when they did not have heuristic worked examples as a scaffold while students with high prior achievement made greater gains with heuristic worked examples. Schwaighofer, et al. (2017) found that there was no significant difference with respect to the order in which the scaffold was administered. However, fading the initial scaffold (i.e. taking the support away after a certain amount of time) had a positive effect on students' dialogic argumentation skills (i.e. social argumentation wherein students come to consensus by agreeing with one another rather than critiquing).

Lavy (2006) explored how MicroWorlds Project Builder, an interactive computer software that allows one to build shapes on a geoboard, aided students in their argumentation on number theory tasks. In her observation of two students, she found students' empirical investigations with the geoboards were essential to their creation of proofs. She also discovered that one student was less skilled in creating verbal arguments, so pointing to the screen acted as a useful warrant for deducing claims . Ugurel, Morali, Karahan, & Boz (2016) similarly used a computerized environment where students could create shapes and figures for proof tasks in Algebra. They sought to understand how an intervention of visual proofs, or proofs by drawing a diagram, could benefit three high school students' argumentative reasoning. They found students both improved in their abilities to create visual proofs over time, and the students found them helpful for conceptually understanding deductive reasoning.

Stoyle & Morris (2017) used blogs as an intervention with fifth-grade students to understand if it helped them create mathematical arguments with fractions and number systems. The 134 students in their study were split into one of three groups (control, blog, and face-to-face). In the blog group, students interacted with each other in groups via a blog so their record of communication was preserved throughout the 22 days of the intervention. The control group did not participate in any interaction with peers while the face-to-face group interacted in person. Via a pre/post-test design, Stoyle and Morris (2017) found the blog group performed significantly better on a test measuring conceptual mathematics knowledge than the other two groups. They hypothesized that writing arguments in a blog space was more beneficial than the other groups because students had to process their thoughts in written format rather than verbal. Cayton-Hodges (2016) similarly hypothesized that writing in a computerized environment would help students' argumentative capacities in Algebra. They hoped to determine what type of student-tutor interactions were most beneficial for students, but their results were mostly inconclusive.

4.3.2 Calculus/Real Analysis

Two studies used a DGE to explore proofs for calculus topics (Caglayan, 2015; Zembat, 2008). Caglayan (2015) introduced GeoGebra to eight mathematics majors in college as a way to help them see limits of functions dynamically rather than from a static point of view. Caglayan interviewed each of the participants for about two hours to determine how they used a dynamic representation of limits to create and argument. He found students created visual proofs using constituents such as hand gestures and cursor gestures. Students came to understand concepts such as limits at a point, right and left hand limits, and limits at infinity because they were able to reason within the dynamic software by checking values and relating them back to algebraic ways of solving a limit. Relatedly, Zembat (2008) used Geometer's Sketchpad as a means for engaging four college students in creating arguments for derivatives. Students in this study worked on one task where they used Geometer's Sketchpad, a TI-83 graphing calculator, and spreadsheet software and another task where they only used paper and pencil. Zembat (2008) found students used multiple types of reasoning while using technology and were able to make connections between algebraic, graphical, and statistical representations. When using only paper and pencil, students only used procedural reasoning making few connections to other representations.

In an undergraduate Real Analysis course, Roy, Inglis, and Alcock (2017) designed an intervention with 49 college students using a tool called *e-Proof,* which is an audio and visual representation of a proof that can be replayed. The e-Proof highlights certain parts of a proof as they are being explained via an audio recording. They hypothesized that students would perform better on proof tasks using e-proofs because it reduces working memory overload. Using an experimental design, they found students that interacted with e-Proof scored lower on a delayed test than the control group who used a textbook signifying the e-Proof group was significantly worse at retaining the information. To understand why their hypothesis did not hold, Roy, et al. (2017) ran further analyses and determined that students did not have to expend the same amount of effort when learning with e-Proofs compared to learning with a textbook. The authors also suggested the audio portion might have interfered with students' abilities to interact with written text. The findings indicate scaffolding too much to aid students in understanding proofs might be detrimental to their long-term retention. Together, the studies reveal favorable outcomes for using DGEs for argumentation in calculus topics, but Roy et al.'s (2017) study provides a cautionary tale for lowering the cognitive demand of tasks.

5. DISCUSSION

This review revealed technology can contribute to enhancing students' capacities, explorations, beliefs, and motivations related to argumentation and proof. This is not surprising given the general consensus amongst scholars that technology enhances mathematical learning (e.g. Heid & Blume, 2008; NCTM, 2000). Still, the current literature reveals several gaps for which future research might examine. The contextual features presented at the beginning of this study revealed most research utilized less than 30 participants and was conducted for less than one week. It is difficult to understand how learners progress in their abilities to create arguments over time while utilizing technology without longitudinal data. Current research mostly takes an observational approach seeking to analyze students' progress or processes over a short period of time. Much of this observational research has been analyzed within this paper, and it makes an important contribution to the field. However, future research might seek to analyze the effects of technology on larger sample sizes over an extended period of time in order to further generalize findings. The contextual features also revealed research on technological supports for proof and argumentation is well-balanced across many geographic regions.

In alignment with other literature reviews, our systematic search and analysis revealed DGEs are a promising support because they allow students to explore dynamic representations of figures to reason and make informed conjectures. Several scholars suggested students might be prone to accept empirical arguments because they can see the truth of a conjecture through multiple manipulations of an object dynamically. This suggests implications for research and practice. Mariotti's (2000) study revealed students can realize actions in a DGE as a theoretical move, but little is known about how students develop theoretical understandings within a technological environment. Researchers might continue to search for supportive actions in making the transition from empirical investigations to deductive proving in DGEs, and practitioners likely need to scaffold this transition.

The review revealed intelligent tutor systems might scaffold students' experiences with technology and proof, but some research warns against over-scaffolding. Like Roy et al. (2017), other scholars should pursue research with an honest examination of how and when to use technology as a scaffold for proof and argumentation. It is possible that over-relying on technology reduces the cognitive demand of a task thereby taking away valuable learning opportunities. Research might specifically examine the extent to which technological scaffolding improves or weakens the cognitive demand of tasks.

Some research experimentally compared technological environments with paper-and-pencil environments, and the literature generally suggests technology enhances students' affective experiences with proof. In general, the research suggests students enjoy utilizing technology more than paper-and-pencil formats, and technology enhances learners' investigations. However, there is no empirical evidence that technology supports students' abilities to create deductive warrants more readily when compared with paper-and-pencil formats. This further suggest the need for future research and practitioners to search for scaffolds which support learners in transitioning from empirical investigations to deductive proofs.

One novelty of our review is that is reports on literature which examines technology as a support for subjects different from Geometry. Our analysis revealed promising technological supports for subjects such as Algebra or Calculus. Collaboration scripts and heuristic worked examples might enhance learners' argumentative abilities, but the research revealed differences in effects based on the abilities of the learner. Future research might examine this relationship and determine when these technological tools are supportive. The analysis also revealed tools such as blogs or GeoGebra in enhancing learners' capacities with proof in Algebra or Calculus. Blogs are promising as a support because they provide a record of communication from past events. Utilizing blogs, practitioners might be able to allow students to reflect on previous knowledge related to proof based on their posts from the past. GeoGebra also provides many opportunities for visualization and conceptual understanding in Algebra and Calculus.

There is a strong need to update research on technological supports for proof and argumentation. Technology is constantly changing, and therefore, previous findings may not be consistent or relevant as new technological tools develop. Many of the studies in this review utilized tools that are no longer widely used within educational settings because new and better tools have become available. In the future, researchers might explore tools such as GeoGebra, Desmos, and graphing calculators which are commonly used in many mathematics classrooms today, otherwise absent from most of this review of the literature since 2000.

Based on this review, scholars ought to see a gap to conduct more research in mathematical content areas other than Geometry since reasoning and proof applies to all content domains (e.g. NCTM, 2000). One might argue that the terms used in this search (i.e. proof and argumentation) excluded literature in other content areas that make important contributions. We acknowledge this as a potential limitation, but it is important to note that these terms are consistent with current curricular recommendations and standards for school mathematics (NCTM, 2000). Therefore, while this limitation may have contributed to the disproportion of studies favoring Geometry, the terms should be well-represented in other content areas as well. Another potential limitation is only reviewing articles written after the year 2000. While this ensured a manageable systematic review process, future literature reviews might seek to understand how technology research related to proof has evolved over time.

With continued attention on proof and argumentation in the K-16 curriculum, it is important that the field continues to theorize and advance research related to teaching and learning proof. With a growing emphasis on using technology regularly in mathematics classrooms, it is likely the research base in this area will grow over the next several years. This review might act as a source for the continued development of technology as a support for proof and argumentation in moving forward in the future.

REFERENCES

References marked with an (*) were articles included in the systematic review.

Arbaugh, F., Smith, M., Boyle, J. D., Stylianides, G. J, & Steele, M. (2019). *We reason and we prove for all mathematics: Building students' critical thinking, grades 6-12*. Thousand Oaks, CA: Corwin Press.

*Baccaglini-Frank, A., Antonini, S., Leung, A., & Mariotti, M. A. (2013). Reasoning by contradiction in dynamic geometry. *PNA, 7*(2), 63-73.

*Caglayan, G. (2015). Math majors' visual proofs in a dynamic environment: the case of limit of a function and the ϵ–δ approach. *International Journal of Mathematical Education in Science and Technology, 46*(6), 797-823.

*Cayton-Hodges, G. A. (2016). Assessing mathematical argumentation through automated conversation. In Wood, M. B., Turner, E. E, Civil, M., & Eli, J. A. (Eds.), *Proceedings of the 38th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Tucson, AZ: The University of Arizona.

*Christou, C., Mousoulides, N., Pittalis, M., & Pitta-Pantazi, D. (2004). Proofs through exploration in dynamic geometry environments. *International Journal of Science and Mathematics Education, 2*(3), 339-352.

*Cobo, P., Fortuny, J. M., Puertas, E., & Richard, P. R. (2007). AgentGeom: a multiagent system for pedagogical support in geometric proof problems. *International Journal of Computers for Mathematical Learning, 12*(1), 57-79.

*Fukawa-Connelly, T., & Silverman, J. (2015). The development of mathematical argumentation in an unmoderated, asynchronous multi-user dynamic geometry environment. *Contemporary Issues in Technology and Teacher Education, 15*(4), 445-488.

Glaser, B., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research.* Chicago: Aldine.

*Gómez-Chacón, I. M., Albaladejo, I. M. R., & López, M. D. M. G. (2016). Zig-zagging in geometrical reasoning in technological collaborative environments: A mathematical working space-framed study concerning cognition and affect. *ZDM, 48*(6), 909-924.

*Guven, B., & Karatas, I. (2009). Students discovering spherical geometry using dynamic geometry software. *International Journal of Mathematical Education in Science and Technology, 40*(3), 331-340.

*Hadas, N., Hershkowitz, R., & Schwarz, B. B. (2000). The role of contradiction and uncertainty in promoting the need to prove in dynamic geometry environments. *Educational Studies in Mathematics, 44*(1-2), 127-150.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.) *Research in collegiatemathematics education III*, (pp. 234-283). Washington DC: Mathematical Association of America

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education, 31*(4), 396-428.

*Healy, L., & Hoyles, C. (2002). Software tools for geometrical problem solving: Potentials and pitfalls. *International Journal of Computers for Mathematical Learning, 6*(3), 235-256.

Heid, M. K. & Blume, G. W. (2008). *Research on technology and the teaching and learning of mathematics: Volume 1: Research syntheses.* Charlotte, NC: Information Age Publishing.

*Hollebrands, K. F., Conner, A., & Smith, R. C. (2010). The nature of arguments provided by college geometry students with access to technology while solving problems. *Journal for Research in Mathematics Education, 41*(4), 324-350.

Hollebrands, K., Laborde, C, & Sträßer, R. (2008). Technology and the learning of geometry at the secondary level. In M. K. Heid & G. W. Blume (Eds.), *Research on technology and the teaching and learning of mathematics: Vol. 1. Research syntheses* (pp. 155-206). Charlotte, NC: Information Age.

*Jones, K. (2000). Providing a foundation for deductive reasoning: Students' interpretations when using dynamic geometry software and their evolving mathematical explanations. *Educational Studies in Mathematics, 44*(1-2), 55-85.

*Kollar, I., Ufer, S., Reichersdorfer, E., Vogel, F., Fischer, F., & Reiss, K. (2014). Effects of collaboration scripts and heuristic worked examples on the acquisition of mathematical argumentation skills of teacher students with different levels of prior achievement. *Learning and Instruction, 32*, 22-36.

Knuth, E. (2002). Proof as a tool for learning mathematics. *Mathematics Teacher, 95*(7), 486–490.

*Lachmy, R., & Koichu, B. (2014). The interplay of empirical and deductive reasoning in proving "if" and "only if" statements in a dynamic geometry environment. *The Journal of Mathematical Behavior, 36*, 150-165.

Lannin, J. K. (2005). Generalization and justification: The challenge of introducing algebraic reasoning through patterning activities. *Mathematical Thinking and learning*, *7*(3), 231-258.

*Lavy, I. (2006). A case study of different types of arguments emerging from explorations in an interactive computerized environment. *The Journal of Mathematical Behavior, 25*(2), 153-169.

*Leung, A., & Lopez-Real, F. (2002). Theorem justification and acquisition in dynamic geometry: A case of proof by contradiction. *International Journal of Computers for Mathematical Learning, 7*(2), 145-165.

Lin, F. L., Yang, K. L., & Chen, C. Y. (2004). The features and relationships of reasoning, proving and understanding proof in number patterns. *International Journal of Science and Mathematics Education*, *2*(2), 227-256.

*Mariotti, M. A. (2000). Introduction to proof: The mediation of a dynamic software environment. *Educational Studies in Mathematics, 44*(1-2), 25-53.

*Mariotti, M. A. (2002). Justifying and proving in the Cabri environment. *International Journal of Computers for Mathematical Learning, 6*(3), 257-281.

*Marrades, R., & Gutiérrez, Á. (2000). Proofs produced by secondary school students learning geometry in a dynamic computer environment. *Educational Studies in Mathematics, 44*(1-2), 87-125.

*Matsuda, N., & VanLehn, K. (2005). Advanced Geometry Tutor: An intelligent tutor that teaches proof-writing with construction. In C.-K. Looi, G. I. McCalla, B. Bredeweg & J. Breuker (Eds.), In *Proceedings of The 12th International Conference on Artificial Intelligence in Education* (pp. 443-450), Amsterdam: IOS Press.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med*, *6*(7), e1000097.

National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics:* Reston, VA: Author.

National Council of Teachers of Mathematics (NCTM). (2014). *Principles to actions: Ensuring mathematical success for all.* Reston, VA: Author.

*Olivero, F., & Robutti, O. (2007). Measuring in dynamic geometry environments as a tool for conjecturing and proving. *International Journal of Computers for Mathematical Learning, 12*(2), 135-156.

*Paneque, J. J., Cobo, P., & Fortuny, J. M. (2017). Intelligent tutoring and the development of argumentative competence. *Technology, Knowledge and Learning, 22*(1), 83-104.

*Roy, S., Inglis, M., & Alcock, L. (2017). Multimedia resources designed to support learning from written proofs: An eye-movement study. *Educational Studies in Mathematics, 96*(2), 249-266.

Schoenfeld, A. H. (1994). What do we know about mathematics curricula? *The Journal of Mathematical Behavior*, *13*(1), 55-80.

*Schwaighofer, M., Vogel, F., Kollar, I., Ufer, S., Strohmaier, A., Terwedow, I., et al. (2017). How to combine collaboration scripts and heuristic worked examples to foster mathematical argumentation–when working memory matters. *International Journal of Computer-Supported Collaborative Learning, 12(3),* 281-305.

Sen, C., & Guler, G. (2015). Examination of Secondary School Seventh Graders' Proof Skills and Proof Schemes. *Universal Journal of Educational Research*, *3*(9), 617-631.

Sinclair, N., Bussi, B., de Villiers, M., Jones, K., Kortenkamp, U., Leung, A., & Owens, K. (2016). Recent research on geometry education: An ICME-13 survey team report. *ZDM Mathematics Education, 48*(5), 691–719.

Sinclair, N., & Robutti, O. (2013). Technology and the role of proof: the case of dynamic geometry. In Alan Bishop, Ken Clement, Christine Keitel, Jeremy Kilpatrick, &

Frederick Leung (Eds.), *Third international handbook of mathematics education* (pp. 571–596). Berlin: Springer.

*Smith, R. C. (2011). A comparison of middle school students' arguments created while working in technological and non-technological environments. In Wiest, L. R., & Lamber, T. (Eds.), *North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 629-637). Reno, NV: University of Nevada, Reno.

*Soldano, C., & Arzarello, F. (2016). Learning with touchscreen devices: game strategies to improve geometric thinking. *Mathematics Education Research Journal, 28*(1), 9-30.

*Stoyle, K. L., & Morris, B. J. (2017). Blogging mathematics: Using technology to support mathematical explanations for learning fractions. *Computers & Education, 111*, 114-127.

Stylianides, A. J., Bieda, K. N., & Morselli, F. (2016). Proof and argumentation in mathematics education research. In A. Gutiérrez, G. C. Leder, & P. Boero (Eds.), *The second handbook of research on the psychology of mathematics education* (pp. 315–351).

Stylianides, G. J., Stylianides, A. J., & Weber, K. (2017). Research on the teaching and learning of proof: Taking stock and moving forward. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. #–#). Reston, VA: National Council of Teachers of Mathematics. Rotterdam: Sense Publishers.

UK Department for Education. (2014). *Mathematics programmes of study: key stages 1 and 2: National curriculum in England.* Retrieved from https://www.gov.uk/government/publications/national-curriculum-in-england-mathematics-programmes-of-study. Accessed 19 December 2018.

*Ugurel, I., Morali, H. S., Karahan, O., & Boz, B. (2016). Mathematically gifted high school students' approaches to developing visual proofs (VP) and preliminary ideas about VP. *International Journal of Education in Mathematics, Science And Technology, 4*(3), 174-197.

*Vogel, F., Kollar, I., Ufer, S., Reichersdorfer, E., Reiss, K., & Fischer, F. (2016). Developing argumentation skills in mathematics through computer-supported collaborative learning: The role of transactivity. *Instructional Science, 44*(5), 477-500.

*Wong, W. K., Yin, S. K., Yang, H. H., & Cheng, Y. H. (2011). Using computer-assisted multiple representations in learning geometry proofs. *Journal of Educational Technology & Society, 14*(3), 43-54.

*Zembat, I. O. (2008). Pre-service teachers' use of different types of mathematical reasoning in paper-and-pencil versus technology-supported environments. *International Journal of Mathematical Education in Science and Technology, 39*(2), 143-160.

*Zengin, Y. (2017). The effects of GeoGebra software on pre-service mathematics teachers' attitudes and views toward proof and proving. *International Journal of Mathematical Education in Science and Technology, 48*(7), 1002-1022.

BIOGRAPHICAL NOTES

Tye Campbell is a PhD student in the department of Curriculum and Instruction at The University of Alabama in Tuscaloosa, Alabama, USA.

Dr. Jeremy Zelkowski is an associate professor in the Department of Curriculum and Instruction at The University of Alabama in Tuscaloosa, Alabama, USA.