Why research on proof-oriented mathematical behavior should attend to the role of particular mathematical content

Paul Christian Dawkins
Northern Illinois University
pdawkins@niu.edu

Shiv Karunakaran
Washington State University
shivk@math.wsu.edu

Because proving characterizes much mathematical practice, it continues to be a prominent focus of mathematics education research. Aspects of proving, such as definition use, example use, and logic, act as subdomains for this area of research. To yield such content-general claims, studies often downplay or try to control for the influence of particular mathematical content (analysis, algebra, number theory etc.) and students' mathematical meanings for this content. In this paper, we consider the possible negative consequences for mathematics education research of adopting such a domain-general characterization of proving behavior. We do so by comparing content-general and content-specific analyses of two proving episodes taken from the prior research of the two authors respectively. We intend to sensitize the research community to the role particular mathematical content can and should play in research on mathematical proving.

Keywords: Proving, mathematical meanings, comparative analyses

Since at least the time of Euclid's geometry, proving has been understood to characterize mathematics as a discipline. Inasmuch as mathematics educators endeavor to engage students in authentic mathematical activity, they have expended much effort to provide students with meaningful proving experiences and document the emergence of proving as a mathematical practice among novices. While we certainly endorse this agenda for instruction and research, we are concerned that framing mathematical proving as a single, domain-general practice may inappropriately downplay the role particular mathematics content plays therein. We observe two trends in the research literature on mathematical proving: 1) making content-independent claims about mathematical proving using data from a particular mathematical context (i.e. analysis, algebra, number theory, geometry) or 2) eliciting student proving behavior in various mathematical contexts (and non-mathematical ones) to yield content-independent findings. In this paper, we consider the possible consequences for research on mathematical proving of downplaying the role of particular mathematical content. We do not at all intend to deny the validity or value of prior research framed in a content-independent manner (some of which we authored), but rather seek to sensitize the community to possible blind spots induced by common lenses applied to research data and to endorse a research agenda focused on the interplay between proving and particular mathematical content.

To portray such blind spots induced by a research lens, this paper presents dual analysis of two episodes taken from prior studies conducted by the two authors respectively. In each case, we compare 1) a content-independent analysis focused on common constructs from proof-oriented mathematics education research – example use, definition use, proof production, logic – with 2) a content-specific analysis focused on explaining students' proving behavior in situ.

Motivating Trends and Questions

It is common to frame both the research questions and findings using these content-independent constructs such that they form informal subdomains of proof-oriented research. One can find numerous examples of studies on proof-oriented mathematical activity that make content-independent claims about

- example use – Alcock & Inglis, 2008; Karunakaran, 2014; Sandefur, Mason, Stylianides, & Watson, 2013,
- definition use – Alcock & Simpson, 2002; Ouvrier-Buffett, 2011,
- proof production – Dawkins, 2012; Raman, Sandefur, Birky, Campbell, & Somers, 2009; Stylianides & Stylianides, 2009,
- logic – Epp, 2003; Selden & Selden, 1995, and
- understanding of proof – Sowder & Harel, 2003; Stylianou, Blanton, & Rotou, 2015.

It is not our goal to critique these studies per se, but rather to sensitize mathematics education researchers to the consequences of consistently investigating proving while downplaying the mathematical meanings that populate the arguments that students produce.

Why do many proof-oriented studies downplay mathematics content? Even if this question had one answer, no available evidence reveals it. Nevertheless, we proffer some possible explanations. One explanation is psychological. Proof's role in mathematics as a discipline and the mathematics education community's emphasis on mathematical process both lead researchers themselves to conceptualize proving in real analysis as one instantiation of a broader phenomenon. Because we as experts can see some uniformity across our broad experiences with proving, we assimilate instances of proving into our more general understanding. A second explanation involves empirical findings. The growing body of evidence of students' difficulties interpreting, producing, and assessing proofs compels mathematics educators to improve upon proof-oriented instruction. Students perceive the transition into proof-oriented courses as a difficult transition, so it seems natural to partition such courses apart from other aspects of the curriculum (though we agree with Reid's, 2011, argument that proving should become and is becoming integrated as a ubiquitous means of mathematical teaching and learning).

A third explanation relates to the analytic process itself. Mathematics educators frequently use localized data to make analytic generalizations (Firestone, 1993) by constructing frameworks and in-depth characterizations of relatively few cases. While such studies rarely make explicit claims to sample-to-population generalizations, it remains unclear how to situate the resulting theory. For instance, Antonini (2003) presented findings suggesting conditions under which students may produce proofs by contradiction, which previous studies reported as challenging. Antonini describes students' exploration of a geometric conjecture involving transversal configurations, but frames his research hypothesis in the following way:

> In task like "given A what can you deduce?" the conjecture can be produced via the analysis of a non-example. The argumentation that justifies the fact that the generated example is a non-example can be re-elaborated and become part of the argumentation of the conjecture. In this case, the argumentation takes an indirect form. (p. 50)

If patterns in students' proof-oriented behavior can be so characterized using content-independent language, when and why should research findings be framed within the content domain at hand (i.e. geometry or the planar geometry of lines)? Antonini's subjects' proofs (and the solution of the given task) depended upon characterizing pairs of lines as intersecting or parallel, which happen to be familiar definitions that are also negations one of another. We posit that this content-specific feature of the task likely contributed to the students' successful use of indirect proof. This raises the question, when and why should researchers emphasize the role of specific mathematical understandings and meanings in framing and explaining research findings on proving? By providing dual analyses (content-independent and content-focused) of two example episodes, this paper sets forth some answers to these questions.

Comparative Analyses

The following sections set forth our two episodes and the dual analysis thereof. The first episode appeared during a sequence of task-based interviews as part of the first author's investigation of student learning of neutral, axiomatic geometry. Episode 1 features two undergraduate mathematics majors trying to prove the equivalence of Euclid's Fifth Postulate (EFP) and Playfair's Parallel Postulate (PPP). Analysis of Episode 1 also appeared in Dawkins (2012). The second episode appeared during a sequence of task-based interviews with expert and novice mathematics students conducted by the second author. Episode 2 features a graduate student in mathematics, designated an expert prover, attempting a novel analysis task about sequences. Analysis of Episode 2 also appeared in Karunakaran (2014). For the sake of brevity and clarity in this theoretical paper, we omit presenting the full methodologies of these studies, which are available in the cited references.

Episode 1: Proving the equivalence of geometric postulates

For reference, the students' statements and diagrams for EFP and PPP appear in Figure 1. As part of a homework assignment prior to the interview, Kirk and Oren had produced a proof of the equivalence of the two postulates using the auxiliary claim we shall call Theorem *, which states "Given two lines cut by a transversal, if the same side interior angles sum is 180, then the two lines do not meet on that side of the transversal." When asked to explain the postulates, the pair found themselves using language from each to explain the other. Oren noted this circularity and attributed it to the statements' mutual implication. Kirk rather explained that the statements "are the same." Oren alternatively explained the postulates' meaning by extending his forearms to represent parallel lines and noting that any amount of rotation from the parallel position would cause the lines to intersect.

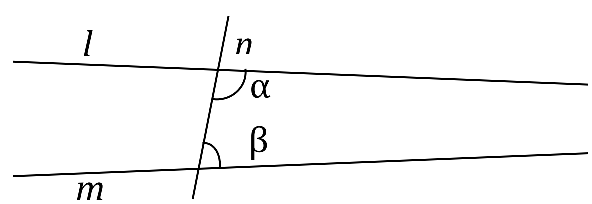
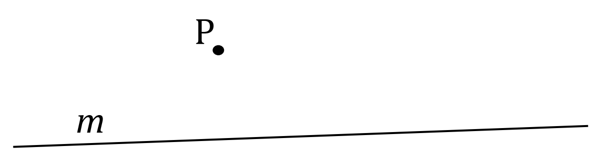
Euclid's Fifth Postulate (EFP): "Given two lines cut by a transversal, if the two interior angles on one side of the transversal sum to less than 180°, then the lines will intersect on that side of the transversal."	
Playfair's Parallel Postulate (PPP): "Given any line and a point not on that line, there exists only one line through the given point that	

Figure 1: Kirk and Oren's statements and diagram for EFP and PPP

The students began the task intending to prove that EFP ⇒ PPP. The students' argument depended upon dividing the line arrangements into three cases, depending upon the angle sum $\alpha + \beta$. They successfully argued, using EFP and Theorem *, that:

- if $\alpha + \beta < 180°$, lines *l* and *m* meet on that side of line *n*,
- if $\alpha + \beta = 180°$, lines *l* and *m* do not meet,
- if $\alpha + \beta > 180°$, lines *l* and *m* meet on the other side of *n*.

Kirk considered this argument sufficient to prove PPP because it guaranteed that there was only one instance in which the lines *l* and *m* are parallel. He said, "Playfair's Postulate basically states that there's only one instance or case where the lines will not meet." Oren disagreed because he was concerned about how the choice of lines through point P (in PPP) corresponded to the angle sums (in EFP). Through their discussion, Kirk also became concerned saying, "It's just hard because Playfair's doesn't include this line n, so you are trying to find a way to go from having this line n to not having this line n in Playfair's." Ultimately, the interviewer invited the students to begin with the diagram for PPP to construct their argument. The pair was able to then use their three cases argument to complete the proof, and Oren correctly identified the need for warrants justifying the construction a transversal line n and guaranteeing that each line l through P corresponded to exactly one angle sum $\alpha + \beta$. Despite their work prior to the interview, Kirk and Oren's proof production took over 40 minutes.

Analysis 1 of Episode 1. The original study in which this episode occurred sought to investigate students' interpretation and use of conditional ("if…then…") statements. The first author used this task because EFP, PPP, and EFP ⇒ PPP can all be understood as conditional statements. Kirk and Oren's initial difficulties in proving EFP ⇒ PPP can be reasonably attributed to the logical structure of their argument, specifically the proof frame (Selden & Selden, 1995). Zandieh, Knapp, and Roh (2008) also reported on students' difficulties with this proof. They attribute this to the fact that students do not adopt a Conditional-Implies-Conditional (CIC) proof frame in which the proof proceeds from the hypotheses of the consequent statement (in this case the point and line arrangement of PPP) to the conclusions of that statement (exactly one parallel through P). Kirk and Oren displayed similar difficulty because they adopted the standard proof frame that begins with hypotheses (EFP) and ends with the conclusion (PPP). Kirk's overall argument could be framed by the valid syllogism "EFP (and Theorem *) ⇒ 3 Cases, 3 Cases ⇒ PPP, therefore EFP ⇒ PPP." However, this argument failed to prove that the conclusions of PPP are entailed in its hypotheses, as the CIC proof does. In Raman et al.'s (2009) language, Kirk understood the *key idea* of the proof (3 Cases argument), but lacked the *technical handle* (the proof frame) to construct a valid proof. Ultimately, the interviewer had to prompt the pair to begin with the diagram from PPP, which implicitly introduced the CIC proof frame. This modification allowed the students to produce a valid and more normative proof.

Analysis 2 of Episode 1. Several aspects of Kirk's behavior in the episode are not explained by the absence of an appropriate proof frame. For instance, why was Kirk convinced by his 3 Cases argument while Oren was not? Also, when Kirk described their intention to prove PPP from EFP, he appeared to metonymize (Zandieh & Knapp, 2006) the two statements by their diagrams. To get from EFP to PPP, one diagram needed to be transformed into the other, which required removing a transversal. We posit that a viable explanation for these phenomena requires attention to the geometric nature of Kirk's reasoning (in a visual-spatial sense). Much like Oren's explanation using his forearms to observe the possible arrangements of two lines, Kirk seemed to interpret the postulates as describing geometric possibilities in a quasi-empirical way. This explains why Kirk metonymized the postulates by their diagrams and said they were "the same" (rather than implied each other): the statements described the same set of geometric possibilities.

Analytically, this account of Kirk's reasoning suggests an alternative syllogistic model: "EFP (and Theorem *) ⇒ Only One Instance, PPP ⇒ Only One Instance, therefore EFP ⇒ PPP." Though this argument is invalid, it reflects Kirk's understanding that the statements are linked because they describe the same possible arrangements of lines. However, each implication in this syllogism is distinct in meaning. His explanation suggested that he viewed Only One Parallel as a paraphrase of PPP rather than a consequence of it. Furthermore, his initial argument did not suggest any directionality to his conclusion since the statements were "the same." Thus, Kirk's empirical reasoning convinced him that the 3 Cases argument proved that EFP ⇒ PPP. Oren, in contrast, seemed to interpret the task of proving in a more conventional hypothetical-deductive manner in which warrants justify inferences that form a chain from hypotheses to conclusions. In short, a researcher imposing a deductive frame on Kirk's reasoning easily misrepresents it.

Episode 2: Proving and disproving conjectures about sequences of real numbers

Upon being asked to validate or refute the mathematical statement given in Figure 2, Zander immediately stated, "So, the first thing that I would do is to see if [the series] obviously doesn't converge." He was asked to further talk about what he aimed to do, and Zander stated that he would search for a counterexample to the statement. That is, he would look for a sequence $\{a_n\}_{n=1}^{\infty}$ of real numbers satisfying the condition that $0 < a_n \leq a_{2n} + a_{2n+1}$, such that the series $\Sigma_{n=1}^{\infty} a_n$ does not converge.

TASK 1

Validate or refute following statement:

Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of real numbers such that $0 < a_n \leq a_{2n} + a_{2n+1}$, $\forall\, n \in \mathbb{Z}$ & $n \geq 1$. Then the series $\sum_{n=1}^{\infty} a_n$ converges.

Figure 2. The statement of the original Task 1 statement as presented to Zander.

Zander quickly generated the valid counterexample sequence $a_n = 1\ \forall\, n$. At this juncture, the interviewer asked Zander to prove a slightly modified version of statement in Task 1. The modified statement read, "Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of real numbers such that $0 < a_n \leq a_{2n} + a_{2n+1}$, $\forall\ n \in \mathbb{Z}\ \&\, n \geq 1$. Then the series $\Sigma_{n=1}^{\infty} a_n$ ***diverges***."

As before, immediately after being given the modified task statement, Zander stated, "Ok. Uh well … right so then I would have to find an example where it converges." The interviewer asked Zander to confirm whether this meant that he was looking for a counterexample to the modified statement, which he did. Also, Zander quickly considered and discarded the use of various tests for convergence and divergence (e.g. ratio test; comparison test) because he anticipated that none of the tests would "guarantee divergence."

Then, Zander recalled an example of a convergent series with which he seemed familiar: the series $\Sigma_{n=1}^{\infty}\left(\frac{1}{2^n}\right)$. He stated his intentions for choosing this example saying, "maybe we can find a way uh to make a sequence where a_n [term from the sequence described in the task] is equal to $\frac{1}{2^n}$ or smaller than or something like that. Cause then that would converge as well." However, he noted that the corresponding sequence does not satisfy the inequality condition $0 < a_n \leq a_{2n} + a_{2n+1}$. To work around this, he attempted to generate a counterexample by modifying the sequence $\left\{\frac{1}{2^n}\right\}_{n=1}^{\infty}$ such that each of the terms repeat using the rule $a_{2n} = \frac{1}{2}a_n$ and $a_{2n+1} = \frac{1}{2}a_n$, and with $a_0 = a_1 = a_2 = a_3 = 1$. At this point he realized that "halving" the terms was the "best–case scenario" in order to satisfy the inequality since "it's sort of the cutoff I mean because if we take it to be any smaller a half, say like a tenth of a tenth and then it no longer fulfills this second inequality."

At this point, he stated that he now believed the modified task statement to be true. Zander then called on the harmonic series to attempt to prove that the modified statement is true, even though the harmonic series does not satisfy the inequality condition. He explained that he would like to show that the terms of the harmonic series (or some variant of it) would be a necessary lower-bound to the corresponding terms of the series in the task and thereby the series in the task would also have to diverge (using the comparison test).

Analysis 1 of Episode 2. The study in which this episode occurred focused on finding similarities and differences between expert and novice's proving behaviors. As such, the original analysis characterized Zander's proving behaviors across the various real analysis tasks provided. Zander used the strategy of searching for a counterexample on this and other tasks. When asked about why he did so, he replied, "Because the counterexample might tell you why it always diverges … or rather the inability to find a counterexample might tell you why it always converges." So, on multiple tasks Zander used this strategy of searching for a counterexample to either successfully find a counterexample invalidating the statement or to gain knowledge about why the statement is valid through the inability to find a counterexample. The interviewer also asked Zander why he called on the series series $\left\{\frac{1}{2^n}\right\}_{n=1}^{\infty}$ and the harmonic series, even though neither one satisfies the inequality condition. He explained that he routinely looked for examples that was relevant to the task context and would provide him with "a picture" or a "prototypical" example that helped him understand the task better.

Thus, this episode supports the general claim that Zander's proving strategy often included searching for counterexamples (regardless of whether he believes one exists), which he perceived useful because he can either successfully find a counterexample or he would gain some insight into why the search for the counterexample is failing and that could tell him why the statement may be true. Furthermore, Zander's work within this episode also supported the claim that he routinely used what he considered "prototypical" examples or visualized "pictures" to gain insight into why a particular claim is true, consistent with previous finding associating visualization and examples with conviction and insight (e.g. Alcock & Simpson, 2004).

Analysis 2 of Episode 2. Even though we can make the content–general claims present in Analysis 1, this may not account for his "expertise" or his relative success on this task. We observe nuances within Zander's search of counterexample and his choice of example series ($\left\{\frac{1}{2^n}\right\}_{n=1}^{\infty}$ and the harmonic series) that provide insights about his use of his content-specific knowledge about series. Throughout the task, Zander paid particular attention to the growth patterns of various series, which can rightly be considered a link between the inequality condition and the convergence of monotone increasing series. When Zander searched for a counterexample for the modified task statement, he called on the series $\left\{\frac{1}{2^n}\right\}_{n=1}^{\infty}$ because he knew this to be a series that converged. However, he noted that this series did not satisfy the inequality condition, but by examining the rate at which the terms of the this sequence decreased, he switched his strategy to find "a way uh to make a sequence where a_n is equal to $\frac{1}{2^n}$ or smaller than or something like that … then that would converge as well." So, Zander deduced that "halving" the terms of the sequence would be the "best–case scenario" since,

> "if we take it to be any smaller a half, say like a tenth of a tenth and then it no longer fulfills this second inequality [and] if we take something that was bigger than a half then that's only more problematic because you're just throwing in bigger numbers into the sequence … I think this if I'm right in saying that this sequence always diverges this actually might be a key to the reason why."

In what ways was this scenario "best?" Zander wanted to find a series that converged, so the added terms must decrease, but the inequality limited the rate at which they decreased. Zander's modified example was his "best" possibility to have a minimal growth rate (so as to converge) while satisfying the inequality condition in the task. It seems that a pivotal reason for Zander beginning to believe that the modified statement is valid is because he noticed that the terms of any sequence that would satisfy the inequality condition would have to have a particular growth rate that was not too fast or not too slow. He called upon the harmonic series (even though it is not a series that satisfies the inequality condition) as a "prototypical" example of a divergent series with a small growth rate, since the sequence of terms added converges to 0. Part of what made Zander's proving successful (his "expertise") was his ability to interpret the conditions in the task as constraints on the growth rate of the series and call upon canonical examples that displayed particular growth behaviors. Both his knowledge and use of the prototypical examples point to his analysis-specific knowledge of series, growth rates, and comparison proof methods.

Discussion and Conclusions

We present dual analyses of these two brief proving episodes to portray the alternative insights gained by content-general analysis (of logic, argumentation, example use, etc.) versus content-specific analysis (of empirical or hypothetical/deductive reasoning, growth rates of sequences and series, etc.). Our two studies reflect common research paradigms within mathematics education: 1) task-based interviews intended to elicit instances of mathematical behavior related to a general topic of interest and 2) comparing and contrasting expert/novice mathematical behavior. While both studies employed grounded theory methods, affording these various analyses, these studies still began with guiding questions and theoretical framings (as no investigation can avoid being, on some level, theory-laden). Regarding Episode 1, it was only after attempts to generally characterize Kirk and Oren's interpretations of conditional statements failed that the author attended to the broader differences between the ways they interpreted the statements and the task at hand, which explain their very different assessments of their proving activity. Regarding Episode 2, the second author designed the study to include tasks in various mathematical contexts, but later refined the study tasks to only include real analysis tasks. While the content-general claims about Zander's proving expertise are supported by Zander's proving practice and his self-reflection, they may also hide the role and value of Zander's extensive experience with real analysis in his interpretation and progress on the task.

As we stated before, our goal is not to deny the value of content-general proof research, but rather to sensitize the mathematics education research community to the liabilities of such a research lens. When and why should researchers attend to the role of particular content in their findings? The first episode suggests that content-general models of student activity such as logic may be broadly applied, but may also be misleading or dishonest to a student's reasoning process. We maintain that the two syllogisms are, in some sense, viable renderings of Kirk's reasoning, but the non-uniqueness of such logical models of his reasoning is troubling. Also, the three "implications" in the latter model of his reasoning are all distinct in meaning and likely gloss over the nature of Kirk's inferences. To hazard an analytic generalization, researchers must be wary applying a content-general model to student reasoning, especially when the chosen model reflect the *researcher's questions* more than the *students' mathematical behavior*.

The second episode suggests that characterizations of "successful proving" or "expertise" must account for the fact that both proving behavior and expertise are highly multi-dimensional. Certainly example use is an important dimension of Zander's proving behavior, as evidenced by his own awareness and explanations thereof. However, the use of such content-general heuristics for further research and instruction necessitate awareness of how example use interacts with other elements of Zander's experience and understanding to afford the behaviors observed in Episode 2. In general, we encourage more research on proving behavior to attend to the role of mathematical meanings (Thompson, 2013). Furthermore, the growing presence of (content-general) introduction to proof courses (Selden, 2012) entails a great need for research on the existence and development of content-general proving behaviors and how they can be fostered within and across mathematical contexts.

References

Alcock, L. & Ingles, M. (2008). Doctoral students' use of examples in evaluating and proving conjectures. *Educational Studies in Mathematics, 69*(2), 111-129.

Alcock. L. & Simpson, A. (2002). Definitions: Dealing with categories mathematically. *For the Learning of Mathematics, 22*(2), 28–34.

Alcock, L. & Simpson, A. (2004). Convergence of sequences and series: Interactions between visual reasoning and learner's beliefs about their own role. *Educational Studies in Mathematics, 57*, 1-32.

Antonini, S. (2003). Non-examples and proof by contradiction. In N. A. Pateman, B. J. Dougherty, & J. Zilliox (Eds.), *Proceedings of the 2003 joint meeting of PME and PMENA* (Vol. 2, pp. 49–55). Honolulu, Hawai'i: University of Hawai'i.

Dawkins, P. C. (2012). Extensions of the semantic/syntactic reasoning framework. *For the Learning of Mathematics, 32*(3), 39–45.

Epp, S. (2003). The role of logic in teaching proof. *The American Mathematical Monthly, 110,* 886-899.

Firestone, W. (1993). Alternative arguments for generalizing from data as applied to qualitative research. *Educational Researcher, 22*, 16–23.

Karunakaran, S. (2014). *Comparing bundles and associated intentions of expert and novice provers during the process of proving* (Doctoral dissertation). The Pennsylvania State University, State College, PA.

Ouvrier-Buffet, C. (2011). A mathematical experience involving defining processes: in-action definitions and zero-definitions. *Educational Studies in Mathematics, 76*, 165-182.

Raman, M., Sandefur, J., Birky, G., Campbell, C., & Somers, K. (2009). "Is that a proof?": Using video to teach and learn how to Prove at The university Level. In *Proceedings of the ICMI Study 19 Conference: Proof and Proving in Mathematics Education* (Vol. 2, pp. 154-159).

Reid, D. (2011). Understanding proof and transforming teaching. In Wiest, L. R., & Lamberg, T. (Eds.) *Proceedings of the 33rd Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 15-30). Reno, NV: University of Nevada, Reno.

Sandefur, J., Mason, J., Stylianides, G., & Watson, A. (2013). Generating and using examples in the proving process. *Educational Studies in Mathematics, 83*(3), 323–340.

Selden, A. (2012). Transitions and proof and proving at tertiary level. In Hanna, G. & De Villiers, M. (Eds.) *Proof and Proving in Mathematics Education: The 19th ICMI Study*, pp. 391-420. Dordrecht, The Netherlands: Springer.

Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics, 29*(2), 123–151.

Sowder, L., & Harel, G. (2003). Case studies of mathematics majors' proof understanding, production, and appreciation. *Canadian Journal of Math, Science & Technology Education*, *3*(2), 251-267.

Stylianides, A. & Stylianides, G. (2009). Proof constructions and evaluations. *Educational Studies in Mathematics, 72*, 237-253.

Stylianou, D. A., Blanton, M. L., & Rotou, O. (2014). Undergraduate students' understanding of proof: Relationships between proof conceptions, beliefs, and classroom experiences with learning proof. *International Journal of Research in Undergraduate Mathematics Education*, *1*(1), 91-134.

Thompson, P. (2013). In the absence of meaning. In K. Leatham (Ed.), *Vital directions for research in mathematics education*. New York: Springer.

Zandieh, M. & Knapp, J. (2006). Exploring the role of metonymy in mathematical understanding and reasoning: The concept of derivative as an example. *The Journal of Mathematical Behavior, 25*, 1-17.

Zandieh, M., Knapp, J. & Roh, K. H. (2008) When students prove state- ments of the form $(P \rightarrow Q) \Rightarrow (R \rightarrow S)$. *Proceedings for the Eleventh SIGMAA on RUME Conference on Research in Undergraduate Mathematics Education*. Retrieved on 1/9/12 from sigmaa.maa.org/rume/crume2008/Proceedings/Proceedings.html.