Student Conceptions of What it Means to Base a Proof on an Informal Argument

Dov Zazkis¹ · Matthew Villanueva²

Published online: 5 August 2016
© Springer International Publishing Switzerland 2016

Abstract In this paper we explore how students construe what it means for an informal argument to be the basis of a proof. We use students' attention as a proxy for their conceptions and document what they pay attention to when assessing whether a proof is based on an informal argument. The data illustrate that some undergraduate mathematics students pay attention to only a limited cross section of the connections that exist between informal arguments and proofs. This indicates that their conceptions of what it means to base a proof on an informal argument involve relatively superficial similarities. Students' attention to only a small cross section of what can be carried forward from an informal argument toward the goal of creating a proof can account for some student difficulties with, and avoidance of, using their own informal arguments as a basis for proof construction.

Keywords Proof · Informal argumentation · Formalization · Semantic/syntactic

Introduction

We explore students' comparisons of informal arguments and formal proofs. Before discussing relevant literature, we clarify some terminology. Our treatment of informal argument and formal proof is similar to Weber and Alcock's (2004, 2009) operationalization of semantic reasoning and syntactic reasoning. Both sets of

✉ Dov Zazkis
zazkis@gmail.com

Matthew Villanueva
mattvill@rci.rutgers.edu

¹ School of Mathematical and Statistical Sciences, Arizona State University, P.O. Box 871804, Tempe, AZ 85287-1804, USA

² Graduate School of Education, Rutgers University, 10 Seminary Place, New Brunswick, NJ 08901, USA

constructs revolve around the observation that proofs are expected to be written in a verbal-symbolic representation system, but may not be generated wholly within that system. Weber and Alcock's constructs are defined relative to this representation system. Producing a proof by working solely within the system is referred to as *syntactic reasoning* and working outside of this system is referred to as *semantic reasoning*. Similarly, we conceptualize a *formal proof* as a deductive argument that establishes the result to be proven and conforms to the norms of the representation system of proof. This is intentionally a characterization of the end product of a proving process and implies nothing about the nature of the reasoning that led to the proof's construction. In contrast, we conceptualize an *informal argument* as a deductive argument that establishes the result to be proven but does not conform to the norms of the representation system of proof. An informal argument can be thought of as a provisional product used to inform formal proof construction. We refer to the use of informal arguments to inform the construction of formal proofs, or more generally the process of using semantic reasoning to inform syntactic reasoning, as *formalization*.

Research relevant to formalization can be partitioned into two overlapping categories. The first category focuses on the role of semantic reasoning in informing proof productions. The second category examines the semantic-to-syntactic formalization process needed to use semantic reasoning to generate a formal proof.

Research falling into the first category has highlighted the important role that semantic reasoning can play in proof generation. Various types of semantic reasoning have been shown to inform proof generation (Gibson 1998; Sandefur et al. 2013; Zazkis et al. 2016). For example, drawing diagrams (Gibson 1998), generating examples (Sandefur et al. 2013), conjecture generation (Boero 1999; Pedemonte 2007), and even the generation of fully formed semantic arguments (Zazkis et al. 2016), have all been shown to inform students' proof generation. Additionally, some researchers have argued that the generation of certain types of proofs is difficult without conceptual insights gained from semantic reasoning (Raman 2003; Weber and Alcock 2004). This first body of work has perpetuated the recommendation that students should use semantic reasoning during proof construction. This recommendation has gained considerable traction in mathematics education (e.g., Garuti et al. 1998; Raman 2003; Weber and Alcock 2004).

A second set of studies has focused on the formalization process itself. This research has provided insights into what types of tasks and mathematical contexts support formalization (Garuti et al. 1998; Pedemonte 2002, 2007). However, this research has also provided evidence that mathematics majors struggle to use semantic reasoning to inform proof generation (Selden and Selden 1995; Alcock and Weber 2010). Selden and Selden (1995) observed that mathematics students were rarely successful in unpacking informally stated mathematical statements and expressing them formally. In their study of 61 students enrolled in a transition to proof course, students successfully formalized only 8.5 % of the statements they were given. Additionally, in their study of 73 mathematics majors, Zazkis et al. (2016) observed that only 3.5 % of student proof attempts involved generating an informal argument during the proving process. Only about half of these informal arguments were successfully formalized. Zazkis et al.'s result suggests that even students who choose to generate informal arguments during proof construction often do not succeed in using these arguments to generate a proof.

These two sets of studies point to a discrepancy between researcher recommendations (that students should generate proofs using informal arguments) and students' behavior and tendencies.

This discrepancy raises the question of what students are attending to when considering informal arguments as the basis of the act of formalization. In order to explore what mathematics majors pay attention to, we studied their assessments of whether a given formal proof is based on a given informal argument. We then compared their assessments to normatively correct assessments. We used our own mathematical expertise and the aforementioned literature to inform what we consider to be normatively correct assessments. Our approach effectively shifts the focus of research from the act of formalization and how it can be supported to students' perceptions of the act of formalization. The student responses provided a window into their conceptions of what it means for an argument to be formalized. Therefore this work posits some of the reasons for the gap between researcher recommendations and student behavior. In turn this work sheds some light on how targeting student conceptions of formalization can narrow this gap.

Theoretical Background

In the next two subsections we discuss what it means for an argument to be the basis of a proof, from our perspective, and introduce several constructs borrowed from cognitive unity research that we repurpose later.

How is a Proof Based on an Informal Argument

Proofs are expected to be written in a verbal-symbolic representation system (but may not be generated wholly within that system). As mentioned earlier, following Weber and Alcock (2004, 2009) we refer to productions that remain wholly within this verbal-symbolic system as *syntactic* and mathematical productions that fall outside of this system as *semantic*. Expanding on our previous definitions, we conceptualize an informal argument as a chain of inferences, at least some of which are semantic, which end in the result to be proven. Similarly, we conceptualize a formal proof as a chain of wholly syntactic inferences that conforms to the norms of the verbal-symbolic system and ends in the result to be proven.

We operationalize a formal proof as *based on* an informal argument if there is a mapping between these two chains of inferences, which has two properties. First, to the extent allowed for by the rules of the verbal-symbolic system, the mapping is meaning preserving. So, as much as possible, inferences in the informal argument map to inferences (or chains of inferences) that have compatible meanings in the formal proof. Second, the mapping preserves logical structure. So corresponding inferences (or chains of inferences) appear in the same order. We use the acronym FBI-judgment, to refer to judgments of whether Formal proofs are Based on Informal arguments. What we mean by these notions will be further elaborated via example when we present the tasks used in this study in [Materials](#) section.

Distance Between an Argument and a Proof

The cognitive unity perspective focuses on the continuum between the production of student generated semantic products, such as unproven conjectures, graphs, examples and informal arguments, and proof production (Boero et al. 1996; Pedemonte 2002, 2008). *Cognitive unity* is said to occur when this continuum exists. That is, cognitive unity occurs when there is a coherent link between students' argumentation activity and their proof generation. Researchers working from the cognitive unity perspective have put substantial effort into examining the continuum between informal arguments/argumentation and formal proofs (e.g., Boero et al. 1996; Garuti et al. 1998; Pedemonte 2002, 2007, 2008).

It is notable that our focus on FBI-judgments, which involve researcher-generated informal arguments and proofs, differs from cognitive unity research, which focuses on argumentation/proofs that are student generated. So in spite of the common focus on formalization, cognitive unity is not directly applicable to our chosen research foci. Even though we cannot adopt the cognitive unity perspective, it is important to acknowledge the parallels between these two bodies of work. Specifically, Pedemonte's (2007) notions of concept distance and structural distance have natural analogs in FBI-judgments. In order to acknowledge these analogs we intentionally borrow part of these constructs' names when generating our own constructs. The *content distance* construct accounts for the differences between how ideas are represented in the proof and informal argument. For example, a student may notice a connection between the semantic (graphical) representation of an increasing function as a function that moves up from left to right and the formal syntactic definition of increasing ($x_1 > x_2$ implies $f(x_1) > f(x_2)$). If a student does not notice such connections between the elements of an informal argument and the corresponding elements of a proof then there is said to be content distance and that student is likely to struggle with formalization. The *structural distance* construct accounts for the layout of the connections between inferences. That is, *structure* refers to what is used to justify each claim in the proof and how these justifications/claims are connected to each other. If the arrangement of these justifications/claims in an argument is substantively different from those in a formal proof then there is said to be structural distance and it is assumed that students will struggle with generating the proof.¹ Both structural distance and content distance play a role in influencing whether the formalization process is successful (Pedemonte 2007).

The Study

We are interested in what it means for a proof to be based on an informal argument from a mathematics major's perspective. We use student attention as a proxy for these conceptions and create a model of what mathematics majors pay attention to when

¹ Pedemonte's (2007) research focuses on specific argument structures, such as abductive arguments, where arrangement of links between inferences is substantively different from those allowed in a proof. However, we interpret her use of the structural distance construct as encompassing all differences in how inferences are arranged.

making informal to formal comparisons and use this model to create a plausible explanation for why these students may have difficulty with formalizing informal arguments.

Participants

This study was conducted at a large state university in the Northeastern United States. The eight participants were pursuing undergraduate degrees in mathematics and were familiar with reading and writing proofs in a variety of mathematical contexts; they had completed a proof based second course in linear algebra, an introduction to proof course, and an introductory analysis course. Participants were intentionally selected to have a range of grades; roughly equal numbers of A, B, and C students were represented. This was used to help ensure our participants had a range of mathematical capabilities.

Procedure

The second author conducted one-on-one clinical interviews with each participant that lasted 90–120 min. During these interviews, the second author presented participants with both informal arguments and proofs and engaged participants in a series of three comparison tasks. The comparison tasks asked students to judge whether particular proofs were based on particular arguments, what they paid attention to, and what they used as evidence when making these judgments. More specifically, we presented participants with triples that consisted of one informal argument and two formal correct proofs, only one of which, from our perspective, was based on the argument. To distinguish between the two proofs in the triple we refer to the proof that is based on the informal argument as the *formalization proof*, and to the one that is not based on the argument as the *distractor proof*.

The informal argument in each triple was presented in the form of a video. Each video lasted approximately 30 s and involved the first author justifying the result semantically with a combination of verbal argumentation, graph generation and gestures. The informal arguments were thus similar to arguments that might be generated by someone working toward producing a proof. Additionally, participants were provided with a transcript of what was said in the video as well as a replication of the diagram produced in each video. The two correct formal proofs in each triple were presented in written form.

At the beginning of each one-on-one interview participants were told that they would be shown triples. They were told that they were to understand and compare the three parts of the triples. They were also told that the given proofs were correct and they did not need to validate correctness. The rest of the interview occurred in two stages. In the first stage of the interview we partitioned each triple into three sets of side-by-side comparisons. For each triple, students were asked to make side-by-side comparisons of what they noticed in terms of similarities and differences between the parts of the triple. Note that during the first stage only the participants' impressions and what they noticed was elicited. We were careful not to use phrases like "based on" during this stage of the interview.

To insure participants understood the triples, triples were shown one at a time. Students watched the video and read the two proofs out loud. They were allowed to re-

watch the video and re-read the proofs as many times as they desired until they felt they understood each part of the triple. Participants were encouraged to ask clarification questions during this process. However, they rarely asked such questions.

The second stage involved revisiting each triple and judging whether each of the proofs in the triple was based on its informal argument. Here students were asked if they had the impression that each of the two proofs in a triple was based on the informal argument, how confident they were in their judgment and what information they used to reach their conclusion. Students were not informed that, from the researchers' point of view only one of the proofs in each triple was a formalization of the informal argument. This made it possible for participants to conclude that neither or both of the proofs in each triple were based on the informal argument.

Analysis

Prior to the collection of interview data the authors discussed the triples in order to generate pre-agreed upon minimum criteria for what a student would need to notice in order to conclude that a particular proof is based on or not based on an informal argument. After the interview data were collected, this minimum standard facilitated the identification of instances where participants made normatively correct FBI-judgments for correct and complete reasons. These standards are discussed in detail in [Materials](#) section.

We were also interested in what students paid attention to when making FBI-judgments. Each interview was transcribed. Constant comparison (Strauss and Corbin 1990) of which aspects of proof students noticed and/or used as evidence when making FBI-judgments allowed for the refinement of categories and facilitated the creation of the model we discuss in the results section. The names of these constructs were later adjusted to reflect parallels between our work and cognitive unity.

Materials

In this section we discuss the specific sets of informal arguments and formal proofs used in our study. Consistent with our definitions in [How is a proof based on an informal argument](#) section the formal proofs are wholly syntactic and the informal arguments rely on informal language and diagrams, and are therefore semantic. Our discussion in this section familiarizes the reader with the study's materials, our interpretations of what it means for a proof to be based on an informal argument and the pre-agreed upon minimum standards used to assess students' informal to formal connection making. We use this standard as a proxy for normatively correct judgments.

To clarify our discussion for the reader, each step in the informal arguments has been labeled with an "I," each step in the formalization of this argument is labeled with an "F" and each step in the distractor proof is labeled with a "D." Additionally, we include subscripts that correspond to the task number to each of the aforementioned labels and break the arguments/proofs into numbered steps. The relationships between the steps in the formalization proofs and steps in the informal arguments discussed in the subsequent sections are instantiations of the mapping we discussed in [How is a proof based on an informal argument](#) section, when we described what it means for a proof to be based on an informal argument.

Task 1: The Integral of Sine Cubed

The first task centered on proofs of the result that the integral of $\sin^3(x)$ over any interval symmetric about zero evaluates to zero. The materials for this task are presented in Fig. 1.

The proof F_1 presented in Fig. 1 starts by justifying that the function, $\sin^3(x)$, is indeed odd by using the fact that $\sin(x)$ is odd as a given (F_1 -1). In the informal argument this fact was implicitly warranted by a graph of $\sin^3(x)$, but not otherwise justified. The proof then uses a series of algebraic manipulations and the fact that $\sin^3(x)$ is odd to show the portion of the integral to the left of the origin is equivalent to the negation of the portion of the integral to the right (F_1 -2 to F_1 -5). This proves that the integral evaluates to zero (F_1 -5 and F_1 -6). Hence, the series of steps from F_1 -3 to F_1 -6 plays the same role in the proof as I_1 -2 and I_1 -3 do in the informal argument. The minimum criteria used for this comparison involves the student noticing that:

- (1) The proof uses the oddness of $\sin^3(x)$ in at least one place (This fact is used in both F_1 -1 and F_1 -5).
- (2) The idea that everything that is “gained” on one side of the y -axis, is “lost” on the other side (I_1 -2) is captured by at least one of the steps (F_1 -2 to F_1 -5).

The distractor proof, D_1 , takes an algebraic approach to justify the result directly. Even though the oddness of $\sin^3(x)$ causes the integral to evaluate to zero in (D_1 -3 and

Task 1: Prove $\int_{-a}^a \sin^3 x = 0$	
Informal argument	Proof F_1 (Formalization of informal argument)
<p>(I_1-1) $\sin^3 x$ is an odd function and the interval is symmetric.</p> <p>(I_1-2) That means that everything you gain on one side, you're going to lose on the other.</p> <p>(I_1-3) So the integral evaluates to zero.</p>	<p>(F_1 - 1) $\begin{aligned} \sin(u) &= -\sin(-u) \\ \sin^3(u) &= -\sin^3(-u) \end{aligned} \quad (\text{cubing both sides}) \quad \left. \vphantom{\begin{aligned} \sin(u) &= -\sin(-u) \\ \sin^3(u) &= -\sin^3(-u) \end{aligned}} \right\} (I_1 - 1)$</p> <p>($F_1$ - 2) Let $-u = x$. So we know that $-du = dx$.</p> <p>(F_1 - 3) $\int_{-a}^a \sin^3(x) dx = \int_0^a \sin^3(x) dx + \int_{-a}^0 \sin^3(x) dx$</p> <p>($F_1$ - 4) $\begin{aligned} &= \int_0^a \sin^3(x) dx + \int_{-a}^0 \sin^3(-u) - du \quad (\text{substituting } u) \\ &= \int_0^a \sin^3(x) dx + \int_{-a}^0 -\sin^3(-u) du \\ &= \int_0^a \sin^3(x) dx + \int_a^0 -\sin^3(u) du \end{aligned} \quad \left. \vphantom{\begin{aligned} &= \int_0^a \sin^3(x) dx + \int_{-a}^0 \sin^3(-u) - du \\ &= \int_0^a \sin^3(x) dx + \int_a^0 -\sin^3(u) du \end{aligned}} \right\} (I_1 - 2)$</p> <p>($F_1$ - 5) $\begin{aligned} &= \int_0^a \sin^3(x) dx + \int_a^0 \sin^3(u) du \quad (\text{because } -\sin^3(-u) = \sin^3(u)) \\ &= \int_0^a \sin^3(x) dx - \int_0^a \sin^3(u) du \quad (\text{flipping the bounds of integration}) \end{aligned} \quad \left. \vphantom{\begin{aligned} &= \int_0^a \sin^3(x) dx + \int_a^0 \sin^3(u) du \\ &= \int_0^a \sin^3(x) dx - \int_0^a \sin^3(u) du \end{aligned}} \right\} (I_1 - 3)$</p> <p>($F_1$ - 6) $= 0$</p>
	Proof D_1 (Distractor)
	<p>(D_1 - 1) $\begin{aligned} \int_{-a}^a \sin^3(x) dx &= \int_{-a}^a \sin(x) \sin^2(x) dx \\ &= \int_{-a}^a \sin(x) (1 - \cos^2(x)) dx \\ &= \int_{-a}^a \sin(x) - \sin(x) \cos^2(x) dx \end{aligned}$</p> <p>($D_1$ - 2) $= \cos(x) - \frac{1}{3} \cos^3(x) \Big _{-a}^a$</p> <p>($D_1$ - 3) $\begin{aligned} &= \cos(a) - \frac{1}{3} \cos^3(a) - (\cos(-a) - \frac{1}{3} \cos^3(-a)) \\ &= [\cos(a) - \cos(-a)] + [\frac{1}{3} \cos^3(a) - \frac{1}{3} \cos^3(-a)] \end{aligned}$</p> <p>($D_1$ - 4) $\begin{aligned} &= [0] + [0] \\ &= 0 \end{aligned} \quad \left. \vphantom{\begin{aligned} &= \cos(a) - \frac{1}{3} \cos^3(a) - (\cos(-a) - \frac{1}{3} \cos^3(-a)) \\ &= [\cos(a) - \cos(-a)] + [\frac{1}{3} \cos^3(a) - \frac{1}{3} \cos^3(-a)] \end{aligned}} \right\} (I_1 - 3)$</p>

Fig. 1 Task 1 - The integral of $\sin^3(x)$

D₁-4) this is not the motivation for the proof. The minimum criteria for making a correct assessment of proof D₁ involves noticing that:

- (1) Oddness (I₁-1) is not explicitly used, and
- (2) The idea that what is “gained” on one side of the y -axis is “lost” on the other (I₁-2) is not a motivation for the proof.

Task 2: Unique Limits of Sequences

The second task was based around proofs of the result that the limit of a sequence, if it exists, is unique. The informal arguments and proofs are presented in Fig. 2.

Unlike F₂, D₂ is not intrinsically an argument by contradiction. The contradiction was artificially added to proof D₂ to make D₂ and F₂ superficially similar. F₂ starts by assuming, toward a contradiction, that the limit is not unique (I₂-1 and F₂-1) and because of this we may choose two limits, L₁ and L₂ (I₂-1 and F₂-2). Although the assumption that L₁ > L₂ (F₂-2) does not explicitly appear in the informal argument, it is implied by the accompanying diagram. Next I₂-2 and I₂-3 argue that the ϵ -neighborhood around L₁ and L₂ can be made small enough to not overlap. In proof F₂, this “not overlapping,” is achieved by choosing ϵ to be exactly half the distance between L₁ and L₂ (F₂-3), and then showing this choice of ϵ places a_n both above and below the midpoint (F₂-4) for all n sufficiently large. Finally both proof F₂ and the

Task 2: If a sequence (a _n) has a limit, it is unique.	
Informal argument	Proof F ₂ (Formalization of informal argument)
(I ₂ -1) If the limit wasn't unique then we could have two limits, say L ₁ and L ₂ .	(F ₂ -1) Toward a contradiction assume the sequence (a _n) does not have a unique limit.
(I ₂ -2) But we control the size of the ϵ -neighborhood around these.	(F ₂ -2) Then there exists L ₁ and L ₂ such that $\lim_{n \rightarrow \infty} a_n = L_1$, $\lim_{n \rightarrow \infty} a_n = L_2$ and $L_1 > L_2$. } (I ₂ -1)
(I ₂ -3) and if we make it small enough these two ϵ -neighborhoods are going to not overlap.	(F ₂ -3) Let $\epsilon = \frac{L_1 - L_2}{2} > 0$. By definition there exists an N such that if $n > N$ then $ L_1 - a_n < \epsilon$ and $ L_2 - a_n < \epsilon$. } (I ₂ -2/3)
(I ₂ -4) So when we get far enough down the sequence we're going to be both in here [pointing an neighborhood around L ₁] and in here [pointing an neighborhood around L ₂].	(F ₂ -4) $ a_n - L_1 < \epsilon \Rightarrow a_n > L_1 - \epsilon = \frac{L_1 + L_2}{2}$ $ a_n - L_2 < \epsilon \Rightarrow a_n < L_2 + \epsilon = \frac{L_1 + L_2}{2}$
(I ₂ -5) But since we can't be in two places at the once we get a contradiction.	(F ₂ -5) But then for $n > N$ we have that $a_n > \frac{L_1 + L_2}{2}$ and that $a_n < \frac{L_1 + L_2}{2}$ which is impossible. } (I ₂ -4/5)
Proof D₂ (Distractor)
	(D ₂ -1) Toward a contradiction assume the sequence (a _n) does not have a unique limit.
	(D ₂ -2) Then there exists L ₁ and L ₂ such that $\lim_{n \rightarrow \infty} a_n = L_1$, $\lim_{n \rightarrow \infty} a_n = L_2$ and $L_1 \neq L_2$.
	(D ₂ -3) By definition for every $\epsilon > 0$ there exists a N such that if $n > N$ then $ L_1 - a_n < \epsilon/2$ and $ L_2 - a_n < \epsilon/2$.
	(D ₂ -4) Consider $n > N$.
	(D ₂ -5) $ L_1 - L_2 = L_1 - a_n + a_n - L_2 $ $\leq L_1 - a_n + a_n - L_2 $ via the triangle inequality $= L_1 - a_n + L_2 - a_n $ $< \epsilon/2 + \epsilon/2$ $= \epsilon$
	(D ₂ -6) $< \epsilon/2 + \epsilon/2$ $= \epsilon$
	(D ₂ -7) So $ L_1 - L_2 < \epsilon$. Since ϵ can be made arbitrarily small $L_1 = L_2$, contradicting the assumption that the limit was not unique.

Fig. 2 Task 2 - Unique limits of sequences

informal argument end by arguing that a_n being in two places at once leads to a contradiction (F_2 -5 and I_2 -4/5). In F_2 this is done formally by arguing that a term in the sequence, a_n , cannot be both above and below the mid point. The minimum criteria used for this comparison involves the student noticing that:

- (1) We are assuming for contradiction that there are two distinct limits (I_2 -1 relates to F_2 -1)
- (2) That we are controlling the size of the epsilon neighborhood, and the choice of epsilon is made so that the neighborhoods do not overlap. (F_2 -3 relates to I_2 -2 and I_2 -3).
- (3) The contradictions are the same, i.e., that eventually, terms in the sequence are in two places at once which is impossible (F_2 -5 relates to I_2 -5).

Proof D_2 demonstrates that any two limits of a sequence can be made arbitrarily close to each other, and thus must be equal. The minimum criteria used for comparing I_2 and D_2 involves the student noticing that:

- (1) The idea of overlapping does not appear in proof D_2 .
- (2) The idea of being in two difference places simultaneously does not appear in proof D_2 .

Task 3: The Derivative of an Even Function is Odd

The third task used in this study is borrowed from Raman (2003). It centers on proofs of the result that the derivative of an even function is an odd function. These are presented in Figs. 2 and 3.

Proof F_3 begins by stating the analytic definition of what it means for a function to be even (F_3 -1). This corresponds to the symmetry-based definition used in the informal argument (I_3 -1). The argument reasons that the symmetry causes corresponding tangents to be reflections of each other. This reflection means that the slope of one tangent is the negative of the slope of its mirror tangent (I_3 -2). Due to the relationship between tangents and derivatives (I_3 -3), this symmetry causes corresponding points on the derivative function to have y-values that are negatives of each other, which is the definition of odd (I_3 -4). In proof F_3 this tangent argument is replaced with the limit definition of derivative (F_3 -2 to F_3 -5). The limit definition uses secant lines, which approach these tangents and can be viewed as a translation of the tangent argument. Algebraic manipulation of the limit definition, which substitutes the definition of odd (F_3 -3), is used to justify that corresponding points on the derivative function have y-values that are negatives of each other. The minimum criteria used for this comparison involves the student noticing that:

- (1) The idea of reflectional symmetry in (I_3 -1) is the basis for the analytic definition used in (F_3 -1)
- (2) The sequence of steps (F_3 -2 to F_3 -5) relates to the steps (I_3 -2 to I_3 -3). Specifically, tangent lines *and their slopes* relate to the limit definition of derivative.
- (3) The conclusions are the same.

Task 3: If a sequence (a_n) has a limit, it is unique.	
Informal argument	Proof F_3 (Formalization of informal argument)
<p>(I_3-1) So even functions have a reflectional symmetry across the y-axis.</p> <p>(I_3-2) So if we take two corresponding tangents, that symmetry means that the tangents are going to have the same magnitude, the same steepness, but they are going to be going in opposite directions.</p> <p>(I_3-3) So that means that when we take the derivative, which corresponds to the slope of these tangents, the derivative is going to have the same magnitude but in opposite directions.</p> <p>(I_3-4) Which means it's going to be odd because odd means things have the same magnitude but go in opposite directions.</p>	<p>(F_3 - 1) If $f(x)$ is even then by definition $f(x) = f(-x)$ } (I_3 - 1)</p> <p>(F_3 - 2) $f'(a) = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$</p> <p>($F_3$ - 3) $= \lim_{x \rightarrow a} \frac{f(-x) - f(-a)}{x - a}$</p> <p>($F_3$ - 4) $= \lim_{x \rightarrow a} \frac{f(-x) - f(-a)}{-1(-x - (-a))}$</p> <p>($F_3$ - 5) $= - \lim_{x \rightarrow a} \frac{f(-x) - f(-a)}{-x - (-a)}$</p> <p>$= -f'(-a)$</p> <p>($F_3$ - 6) So $f'(x)$ is an odd function. } (I_3 - 4)</p>
	<p>Proof D_3 (Distractor)</p> <p>(D_3 - 1) If $f(x)$ is even then by definition $f(x) = f(-x)$. } (I_3 - 1)</p> <p>(D_3 - 2) $f(x) = f(-x)$ $f'(x) = f'(-x) * (-1)$ because of the chain rule $f'(x) = -f'(-x)$</p> <p>(D_3 - 3) So $f'(x)$ is an odd function. } (I_3 - 4)</p>

Fig. 3 Task 3 - The derivative of an even function is odd

The distractor proof, proof D_3 , has the same beginning and ending as proof F_3 . However, instead of translating the symmetry argument, it uses a well-known theorem, the chain rule. There is nothing in the informal argument that points to the use of this theorem. The minimum criteria for this proof involve noticing that:

- (1) No ideas directly related to slope or tangents show up in the proof.
- (2) There is nothing in I_3 that suggests the use of the chain rule.

It is important to mention that Raman (2003) herself identified proof D_3 as based on a procedural idea separate from I_3 and identified proof F_3 as based on the idea illustrated in I_3 . Thus, our interpretations of this triple are compatible with Raman's.

Results

The results are partitioned into two sections. [Relating informal arguments and proof](#) section discusses a top-level view of the connections the participants made relative to the pre-agreed standards discussed in the materials section. This addresses the frequency of participants' normatively correct judgments of the tasks relative to our criteria. [A model of what students pay attention to when making FBI-judgments](#) section posits a model of what students pay attention to when making FBI-judgments. This model is illustrated with specific examples of students' work.

Table 1 Number of normatively correct FBI-judgments relative to the minimum standard

	Task 1	Task 2	Task 3
Met criteria for D-I judgment	7/8	5/8	2/8
Met criteria for F-I judgment	8/8	3/8	1/8
Met both F-I and D-I criteria	7/8	2/8	0/8

Relating Informal Arguments and Proof

In Table 1 we present a top-level view of the data by documenting how many participants met our minimum standard and hence, from our perspective, made normatively correct FBI- judgments for normatively correct reasons.

As can be seen from Table 1, the first task was relatively unproblematic for the students in this study. Only one student, S4, did not meet our pre-agreed upon minimum standard. He was able to make the correct assessments regarding which proofs were based on the informal argument (I_1), however, he did not make appropriate connections when comparing the distractor (D_1) and the informal argument (I_1). With respect to the other two tasks students often did not meet our minimum standard. Only 2 of the eight students met our minimum standard on both parts of task 2 and none of our students met our minimum standard on both parts of task 3.²

These data point to mathematics majors finding some informal-to-formal comparison tasks difficult. This difficulty with recognizing formalization is consistent with students' difficulties formalizing (e.g. Selden and Selden 1995; Zazkis et al. 2016). We believe these difficulties are linked; the ability to recognize the final product of formalization, a proof, as based on an informal argument is necessary but not sufficient for translating that informal argument into a formal proof. Without this ability a student cannot recognize when the formalization process is complete and thus cannot effectively formalize.

A Model of What Students Pay Attention to When Making FBI-Judgments

In our analysis we used the constant comparative method (Strauss and Corbin 1990) to identify four different aspects of arguments/proofs that students focused on when making FBI-comparisons. Two of these are similar to Pedemonte's (2007) structural distance and content distance constructs. So we choose to retain "structural" and "content" as part of the names of our constructs. However, we use "foci" in place of "distance" to emphasize that the focus of Pedemonte's research on translation processes is different from our own focus on FBI-comparisons. Pedemonte's research examined proving and used structural distance and content distance to account for student difficulties with proof generation. Instead, we examine comparisons between ready-made informal arguments in proofs and use student attention to only a subset of the connections between these objects to account for difficulties with FBI-comparisons. The four foci of comparison are described below:

² The tasks were presented to students in the order 1,3,2 and then renumbered during the analysis process so the task number corresponded to the level of student success.

- (1) *Structural foci* involve noticing similarities and differences in how inferences follow from one another (i.e., structure of arguments/proofs). In a broad sense, noticing structural foci can be seen as an attempt to evaluate what kind of relationship exists between the arrangement of links between and within inferences in an informal argument and the arrangement of links between and within inferences in a formal proof. We conceptualize this as an adaptation of Pedemonte's (2007) notion of "structural distance" to a FBI- comparison context. Both our construct and hers center on the role that preservation or lack of preservation of structure play in formalization.
- (2) *Content foci* involve noticing which specific elements (i.e., inferences, assumptions, data and claims) are present or not present within both an argument and proof. This can be seen as an attempt to evaluate the relationship between the content of an informal argument and the content of a formal proof. This focus can be seen as an adaptation of Pedemonte's (2007) notion of "content distance" to an FBI- comparison context.
- (3) *Methodological foci* involve noticing the proof method used (e.g., contradiction, contrapositive, induction, construction, etc.) as well as the role this method plays in the proof.
- (4) *Holistic foci* involve noticing similarities and differences in terms of goals, style purpose or overarching mathematical idea. These comparisons focus on proofs and arguments as a whole and overlook specific structural, content and methodological details.

In what follows we show that prioritizing one of the four foci of comparison in lieu of others has detrimental effects on students' ability to make FBI-comparisons. These foci thus influence how students interpret what it means for a proof to be based on an argument. It is important to note that these foci are not necessarily static. Some students shift foci when moving to a different task and others may remain stable with respect to what they pay attention to when making informal to formal comparisons.

Additionally, we interpret instances where a student makes a confident FBI-judgment while attending to only one of the foci discussed above as an indication that the student has a "based on means share a single (or small number of) common attribute(s)"-conception as opposed to "based on means the proof is an evolution of the informal argument"-conception. We refer to the former as an underdeveloped conception.

Structural Foci

We begin by discussing structural foci. We illustrate structural foci with the work of S4 on Task 3. Below is S4's assessment of whether proof F_3 is based on I_3 :

S4: Is $[F_3]$ based on it? [...] I definitely know it's based on it...

Int: Okay, so it's based on it.

S4: Yeah [...]

Int: Okay, so what ideas show up that made you say that? What connections are you making when you say that?

S4: Well, you're going from even to odd, and one of the first thing is corresponding tangents. And how do you do that? You take the derivative. You're using the definition of derivative right away on an even function. So... maybe you're not using the concept of the tangent line right away, but you're definitely, but you definitely know that you're taking the derivative and that it has to be negative.

Notice that in the above excerpt when comparing I_3 and F_3 S4 indicates that he does not see the tangent idea from I_3 appearing in the proof. In [Materials](#) section, when discussing this triple, we concluded that the connection between the tangent argument in I_3 and the limit argument in F_3 was central to why I_3 is based on F_3 . S4 overtly mentioned that he did not see this connection. His assessment of why I_3 is based on F_3 centers on the fact that the derivative is used to link the definition of even function with the definition of odd function. This is a structural comparison rather than a content comparison because he sees no overt link between the limit definition and the idea of tangent, yet he chooses to conclude F_3 is based on I_3 . That is, he focuses on the existence of a link not the reasons for the link. He is confident in his conclusion as evidenced by his use of the word “definitely.” This confidence occurs in spite of his noticing only a subset of possible connections, pointing to his underdeveloped conception of based on.

Now we look at S4's assessment of whether proof D_3 is based on I_3 :

S4: This is definitely based on it.

Int: So this is based on it again

S4: Yeah

Int: Okay, uh... why? Why do you say it's based on it? What connections do you make between?

S4: Well this one... without even using the definition of derivative, you're just taking the derivative, so... that's... the derivative you're taking the tangent and you're finding that it's a negative tangent to the other one, and you get... just by taking that tangent, that... derivative of the first one is an odd function, which is pretty much step-by-step, going through this (informal argument)

As with the previous excerpt S4 appeals to the parallels in structure between I_3 and D_3 to conclude that D_3 is based on I_3 . He does this with a similar level of confidence to his previously discussed excerpt, again using the word “definitely”. He sees unsubstantiated content similarities between taking the derivative using the chain rule and the tangents in I_3 . However, in spite of

noticing these similarities in D_3 and not I_3 he makes the same assessment for both. Thus the existence or non-existence of content similarities does not appear to sway his FBI-judgment and he bases his conclusions on structure alone.

In both of the above excerpts, when comparing the informal argument, which also uses the idea of derivatives to bridge even with odd, S4 determined that both proofs were based on the informal argument. This conflicts with what we labeled a normatively correct interpretation. Since both the proofs and the informal argument use derivative (I_3 -3) to bridge the idea of even (I_3 -1) with the idea of odd (I_3 -4), this determination is consistent with his prioritization of structure. In other words, since he does not prioritize the content connections between tangents and limits, he does not distinguish between F_3 and D_3 in a meaningful way.

Since any proof of the result in Task 3 must use derivative to link even and odd functions, S4's conception of what it means for a proof to be based on an argument leads to the conclusion that *any* proof of the result is based on the informal argument. This points to S4's structural foci interfering with his ability to make normatively correct assessments of the similarity between proofs and informal arguments. Further, it provides evidence that prioritizing structure in lieu of content may lead to normatively incorrect FBI-judgments.

Content Foci

Next we discuss content foci. This is characterized by students paying attention to the assumptions and inferences within a proof/argument, but largely overlooking the roles and structure of these elements within a proof. Here the focus is localized to specific elements in the proof. In other words, the details of a proof are examined, but the bigger picture is back grounded or wholly ignored. We illustrate this focus with an excerpt of S8's work on Task 2. In the excerpt below S8 is asked to compare D_2 and F_2 :

Int: Any... any other differences you can see? Looking at the proofs again?

S8: Umm... L_1 is defined to just not be equal to L_2 in proof D_2 and in proof F_2 , they say L_1 is greater than L_2 .

Int: I, I guess, ... do you consider those significant differences? The ones that you mentioned?

S8: Yea, yea definitely. Those are significant differences.

The assumption that $L_1 > L_2$ is not explicitly made in I_2 , however, it is implied by the accompanying diagram. Following the above excerpt S8 proceeded to use the fact that the $L_1 > L_2$ assumption is part of proof F_2 and not D_2 as a justification for why the distractor proof is based on the informal argument while the formalization proof is not. In focusing on a particular piece of content, which is present in only one of the proofs, he overlooks the bigger picture and, consequently, makes an FBI-judgment that is

inconsistent with the normatively correct interpretation. Hence, we contend that prioritizing content in lieu of structure is also insufficient for making normatively correct FBI-judgments.

S8's assessment is consistent with his content foci. He is looking for specific inferences that are present in the proofs in order to compare them to the informal argument. Thus, his conception of what it means for a formal proof to be based on an informal argument involves the formal proof using similar assumptions and similar inferences to the informal argument. Within this conception, a difference in assumptions used is sufficient evidence that a proof is not based on an informal argument. S8 makes a confident assessment that a "definitely significant" gap between argument and proof exist based on his noticing a single difference. This points to S8's underdeveloped conception.

Methodological Foci

It is useful to note that our anticipation of methodological foci influenced our task design. Methodological foci were the motivation for making proof D_2 artificially a proof by contradiction. If we had not artificially made D_2 a proof by contradiction students may have concluded that D_2 was not based on I_2 solely based on the fact that I_2 is a proof by contradiction without working to make other connections. Thus, our task design intentionally discouraged surface level methodological foci. However, one participant did notice this feature of D_2 :

S7: Wait, what? ... There is no point in this [D_2] being a proof by contradiction. That is completely redundant I could have just crossed this out here, "Assume it does not have a unique limit." You can cross that out.

Our task design limited opportunities for superficial methodology based assessments but left the tasks open to deeper assessments, like the one in the excerpt above. Since only one of the participants noticed this feature of D_2 we argue that artificially adding or removing particular methodologies from proofs has the potential to lead students to make incorrect FBI-judgments. That is, if we used a direct proof as a second distractor in place of F_2 , we anticipate that the majority of students in our study would have incorrectly used "only one of these is a proof by contradiction" as a justification for why D_2 was based on I_2 . This highlights the limitations of methodological foci, since over prioritizing methodology may lead to overlooking both structure and content.

Holistic Foci

The final type of foci we discuss involves examination of holistic traits. The word trait here is construed broadly and may include attribute such as elegance, efficiency, style, pedagogical purpose or overarching idea. In short, this is intended to capture any treatment of a proof as more than the sum of its parts. Proofs have purposes and can be qualitatively compared to both each other and to the general genre of proof writing.

We begin by discussing the work of S1 on Task 1. The excerpt below takes place after S1 reads proof D_1 (He has already read proof F_1).

S1: I feel like D_1 was kind of lamer than the other one

Int: Lamer?

S1: This one $[F_1]$ was a little prettier, it was... I mean over here, we had uh, we were using that it was... this $[D_1]$ felt very... brute force

S1 goes beyond interacting with these proofs as proofs and treats the two proofs in task 1 as aesthetic entities. He expresses the belief that elegant proofs are more desirable than brute force proofs and judges D_1 as less desirable. Later in his interview, when he was asked to compare F_1 and D_1 , he discussed the two proofs relative to the genre of proof as a whole.

S1: Okay, what do they have in common? Clearly they have the goal in common, but the guy on the left, proof D_1 , proof D_1 felt more like uh... I don't really know if there's an actual distinction in the math world between a "prove something" and a "show something," but if there was, this $[D_1]$ definitely feels like, you know, just show that it's 0. But this $[F_1]$ was like a really... this felt like it had more behind it here... whereas this $[D_1]$ was like, let's just evaluate it and see where that takes us. Okay? Which is fair, you know? It just doesn't give you any insight into why that's the case.

S1 expresses the belief that proof D_1 , does not provide any mathematical insight regarding why the result holds. It is simply an exercise in implementing well-established calculation techniques. Implicitly, he expresses that he often looks for what insights he can gain from presented proofs, in this case he did not find any.

However, students cannot effectively make FBI-judgments by focusing on holistic attributes of a proof in lieu of other attributes. For example, there may be multiple elegant arguments that justify a result. Thus, elegance alone is insufficient for students to make consistently correct FBI-comparisons.

Multi-focus Comparisons

In the previous four subsections we argued that prioritizing only one of the foci of comparison in lieu of others was insufficient for making normatively correct FBI-judgments. In this section we illustrate what comparisons that attend to all four foci look like and how they may yield normatively correct FBI-judgments. To clarify we are not claiming that attention to all four foci is guaranteed to yield normatively correct judgments. Rather, we are arguing that balancing one's attention to these foci greatly increases the likelihood that a student consistently generates correct FBI-judgments.

Below we examine S7's work on Task 2, specifically his reaction when looking at proof F_2 for the first time.

Int: General impressions.

S7: F_2 is just literally the proof version of I_2 Uh, so the idea behind this is that, okay if we are trying to show that this sequence has a unique limit, which we

want to show that it can't have two limits. So we suppose there is two limits, basic proof by contradiction. So both are proofs by contradiction. And the contradiction occurs when epsilon is small enough. Here they show it intuitively but it's pretty clear from the picture that what they use was a number that's less than half way in between. Here $[F_2]$ is that function, the average. So once we have the. It's not the average it's close enough so that it doesn't even reach the average. And that way the two have no overlap. And then by definition of sequence it should eventually get far enough that it's in this region always and once you get far enough it's in this region always but then it's therefore always in both these regions once it passes that specific end that we defined. And that's the contradiction. Which is what they said here. When we get small enough down we can't be in both but it has to be in both.

It is important to emphasize that S7 realizes that F_2 is based on I_2 before he is asked to make any kind of comparison. The above was simply his initial response. The part of the interview where he was specifically asked FBI-questions occurred 30 min later. Also, he immediately jumped into the comparison when he stated that both the proof and the informal argument have the same idea behind them (holistic foci). He then shifted to discussing how this idea manifested itself in terms of structure of both I_2 and F_2 (structural foci). He notices that both the proof and informal argument are necessarily arguments by contradiction with the contradiction in both cases being that you cannot be in two places at once (methodological foci). This is then related to the specifics of the proof, with being both above and below the midpoint of L_1 and L_2 corresponding to being in two places at once in the informal argument (content foci). S7 makes all four types of comparisons, does this without any specific prompting to make a comparison and relates the four types foci in his discussion.

We believe that the fact that S7 saw the relationship between I_2 and F_2 before he was asked to compare them to be important. Mathematics is often discussed metaphorically as a natural language (e.g., Downs and Mamona-Downs 2005). Here S7 recognized that the informal argument, which is expressed using semantic language and semantic representations, and the proof, which is expressed using syntactic language were metaphorically telling the same story. This is akin to being shown two paragraphs that tell the same story in two different languages, both of which one is fluent in. The fact that the same story has been presented twice, as well as the multitude of parallels between its two presentations, is salient even without being asked to compare the two paragraphs.

On the other hand, if one is learning a foreign language and is asked if both paragraphs tell the same story, the comparison is very different. The rich connections between the two tellings of the story are not immediately clear. The comparison becomes an exercise in finding whatever similarities and differences can be readily found and using these as evidence for making a decision. The types of similarities and differences noticed are likely limited, perhaps to the point where only one similarity/difference is noticed. This noticing of only a single type of connection is analogous to what we observed students doing when they prioritized one of the foci over others.

Discussion

This paper contributes to the literature on proof and proving theoretically and methodologically. First, from the perspective of theory, this paper introduced a four-part model of the aspects of arguments/proofs students focus on when attempting to determine whether a particular proof is based on a particular argument. The components of this model are structural foci, content foci, methodological foci and holistic foci. We illustrated that when students focus on one of these in lieu of others they were prone to normatively incorrect FBI-judgments. Further, we argued that comparisons that incorporated multiple types of connections and acknowledged the interrelationships between these were beneficial for making normatively correct FBI-judgments. This variation of types of connections was operationalized via our model as the foci of those connections. The instances where a student prioritized a single focus at the exclusion of others resulted in a normatively incorrect judgments about whether an informal argument is the basis of a proof. We interpret instances where students make incomplete assessments with high confidence as an indication that students conceive of based on as synonymous with “share a single (or small number of) common attribute(s)” as opposed to “the proof is an evolution of the informal argument.” This conception at least partially accounts for difficulties students have with generating proofs based on informal arguments (e.g., Zazkis et al. 2016) and in understanding the connections between informal arguments and proofs presented in lecture (e.g., Lew et al. 2016).

Our methodological contribution is the use of triples of informal arguments, formal proofs, and distractor proofs as a research tool for those interested in research on the connections between informal arguments and formal proofs. Examining how students compare and contrast ready-made informal arguments and formal proofs provided insights regarding what they notice when making FBI-judgments. In turn, what students’ noticed during FBI-judgments provides a lens into how they conceptualize formalization in general and informs possible reasons for students’ struggles with formalizing their own informal arguments. This method reveals that students’ conceptions often only encompass a part of the connections that exist between informal arguments and formal proofs. The triples method used with different sets of triples has the potential to reveal more about how students conceptualize formalization. We plan to use this methodology in subsequent studies.

These observations regarding students’ conceptions point to an additional avenue through which instructors can help facilitate their students’ formalization abilities—bolstering their ability to perceive connections between informal arguments and proofs. That is, FBI-judgments may be used in class as a means for directing students’ attention to the types of connections that exist between informal arguments and proof (ie. holistic, methodological, structural and content connections). Further research is needed to determine if such activities benefit students’ ability to effectively utilize informal arguments during their proving.

Finally, our presentation of informal arguments in video form, with accompanying, gestures, diagrams and imprecise language expands the treatment of what it means for an argument to be informal beyond our operationalization of this construct. This approach specifically treats informal arguments as akin to arguments that students generate while working toward producing proofs, rather than general productions that contain elements not typically allowed in formal proofs.

This study also sheds light on a number of important future directions for research. There was a disparity between student successes on Task 1 in comparison to other tasks. What factors make some FBI-tasks more difficult than others is an interesting question that requires a larger and more diverse set of tasks than the three discussed here. Additionally, a notable constraint of this study is that all the FBI-tasks were in a calculus/analysis context. Pedemonte (2008) observed that some mathematical contexts were more conducive to formalization than others. Whether the same holds true for FBI-comparisons remains an open question.

There are also a host of other studies that have shed light on formalization (e.g., Pedemonte 2007; Selden and Selden 1995; Weber and Alcock 2004; Zazkis et al. 2016). Taken as a whole, these studies have created an impetus for the design of theoretically informed instruction that targets students' formalization. Studying the implementation of this type of instruction at an undergraduate level is an important emerging direction for formalization related research (e.g., Rasmussen and Marrongelle 2006).

References

- Alcock, L., & Weber, K. (2010). Undergraduates' example use in proof production: purposes and effectiveness. *Investigations in Mathematical Learning*, 3(1), 1–22.
- Boero, P. (1999). Argumentation and mathematical proof: a complex, productive, unavoidable relationship in mathematics and mathematics education. *International Newsletter on the Teaching and Learning of Mathematical Proof*.
- Boero, P., Garuti, R., Lemut, E., & Mariotti, M. (1996). Challenging the traditional school approach to the theorems: A hypothesis about the cognitive unity of theorems. In L. Puig, & A. Gutierrez (Eds.), *Proceedings of the 20th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 113–120). Valencia, Spain: Universitat de Valencia.
- Downs, M. L., & Mamona-Downs, J. (2005). The proof language as a regulator of rigor in proof, and its effect on student behavior. In M. Bosch (Ed.), *Proceedings of CERME 4, Working Group 14: Advanced Mathematical Thinking* (pp. 1748–1757). Sant Feliu de Gix, Spain: FUNDEMI JQS - Universitat Ramon Llull.
- Garuti, R., Boero, P., & Lemut, E. (1998). Cognitive unity of theorems and difficulty of proof. In A. Olivier, & K. Newstead (Eds.), *Proceedings of the 22nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 345–352). Stellenbosh, South Africa.
- Gibson, D. (1998). Students' use of diagrams to develop proofs in an introductory real analysis. *Research in Collegiate Mathematics Education*, 2, 284–307.
- Lew, K., Fukawa-Connelly, T. P., Mejía-Ramos, J. P., & Weber, K. (2016). Lectures in advanced mathematics: Why students might not understand what the mathematics professor is trying to convey. *Journal for Research in Mathematics Education*, 47(2), 162–198.
- Pedemonte, B. (2002). Relation between argumentation and proof in mathematics: cognitive unity or break. In J. Novotná (Ed.), *Proceedings of the 2nd Conference of the European Society for Research in Mathematics Education* (Vol. 2, pp. 70–80). Prague: Czech Republic.
- Pedemonte, B. (2007). How can the relationship between argumentation and proof be analysed? *Educational Studies in Mathematics*, 66(1), 23–41.
- Pedemonte, B. (2008). Argumentation and algebraic proof. *ZDM*, 40(3), 385–400.
- Raman, M. (2003). Key ideas: what are they and how can they help us understand how people view proof? *Educational Studies in Mathematics*, 52, 319–325.
- Rasmussen, C., & Marrongelle, K. (2006). Pedagogical content tools: integrating student reasoning and mathematics in instruction. *Journal for Research in Mathematics Education*, 388–420.
- Sandefur, J., Mason, J., Stylianides, G. J., & Watson, A. (2013). Generating and using examples in the proving process. *Educational Studies in Mathematics*, 83(3), 323–340.
- Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics*, 29(2), 123–151.

- Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage.
- Weber, K., & Alcock, L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics*, 56(2–3), 209–234.
- Weber, K., & Alcock, L. J. (2009). Proof in advanced mathematics classes: Semantic and syntactic reasoning in the representation system of proof. In D. A. Stylianou, M. L. Blanton, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. 323–338). New York: Routledge.
- Zazkis, D., Weber, K., & Mejía-Ramos, J. P. (2016). Bridging the gap between graphical arguments and verbal-symbolic proofs in a real analysis context. *Educational Studies in Mathematics*. doi:[10.1007/s10649-016-9698-3](https://doi.org/10.1007/s10649-016-9698-3).