# PERSONAL INFERENCES AS WARRANTS OF UNDERGRADUATE STUDENTS' ARGUMENTS IN CALCULUS CONTEXTS

Kyeong Hah Roh
Arizona State University
khroh@asu.edu

Erika David Parr
Rhodes College
parre@rhodes.edu

Derek Eckman
Arizona State University
dceckman@asu.edu

Morgan Sellers
Colorado Mesa University
msellers@coloradomesa.edu

*The purpose of this paper is to highlight issues related to students' personal inferences that arise when students verbally explain their justification for calculus statements. We conducted clinical interviews with three undergraduate students who had taken first-semester calculus but had not yet been exposed to formal proof writing activities through undergraduate mathematics courses. We analyzed these students' verbal justification of four statements, each of which described a different relationship between two quantified variables in calculus contexts. In this paper, we document students' personal inferences that were evoked in their justifications and discuss how they are similar to or different from the logical inferences that have been accepted and practiced by the mathematics community for mathematical arguments.*

Keywords: Calculus, reasoning and arguments, inferences, Toulmin's model of argumentation.

Reasoning and arguments are foundational aspects of mathematics. NCTM (2000) and Common Core State Standards (2010) have emphasized teaching students in all age groups to develop, evaluate, and use various types of reasoning and methods of mathematical arguments. On the other hand, research in undergraduate mathematics education has reported that students frequently make their arguments on empirical bases (Shipman & Shipman, 2013). Students often believed that examining a few cases would be sufficient as a justification in mathematics (Harel & Sowder, 1998) or stated that both example-based proofs and formal deductive proofs are equally valid (Martin & Harel, 1989). In addition, when reading an argument, students often accept it by ignoring, or without attending to, inferences inherent in the argument (Alcock & Weber, 2005). Dawkins and Zazkis (2021) reported that some students with no experience in mathematical proof construction did not envision proof as an unbroken chain of implications leading from the first line to the last line of an argument. Stylianides et al.'s (2004) study also documented that even undergraduate students who studied mathematics did not accept or use the equivalence between a conditional statement and its contrapositive. A similar phenomenon was also reported in Dawkins et al.'s (2021) study in which students did not conceive of an inference from the contrapositive equivalence in reading an argument and believed that a proof of the contrapositive of a conditional statement would not prove the original conditional statement. Instead, these students accepted an argument directly proving the converse of a conditional statement as a proof of the original conditional statement.

Likewise, we see some gap between students' personal inferences, as utilized in their proof-writing or validating others' arguments, and logical inferences, as accepted and practiced by the mathematics community. In this paper, we examine if such a gap still exists when students verbally express their arguments. Focusing on three undergraduate students who had not yet taken proof-oriented mathematics courses at university levels, we address the following research

question: *What are the types of personal inferences that students naturally use, and how do they use their personal inferences as warrants to support their claims in calculus contexts*?

## Theoretical Perspective

By *personal inference,* we mean an act or process that an individual draws to make a claim from things that are known to him or her. We consider students' personal inferences through the lens of radical constructivism (Glasersfeld, 1995). From this perspective, we assume that students' inferences consist of sets of action schemes which are built and refined through their experience. Students' inferences are personal, in this view, as their unique experiences have shaped the ways they support arguments. Even if students have not yet experienced reading or constructing formal mathematical proofs, they would have made inferences for their arguments through their experience. These personal inferences may or may not align with what the mathematics community considers conventional valid logical inference rules such as modus ponens, modus tollens, or various syllogisms. Additionally, from this perspective, we as researchers are limited to what we can access about students' personal inferences. At best, we can only hope to model students' personal inferences. Further, we acknowledge that the tasks used in this study may have evoked only some of the ways in which students may use personal inference to support their claims when making arguments. In other words, we do not intend to document all of a student's personal inferences. Rather, in this paper, we highlight the students' personal inferences that were *evoked* by the context and task.

We also adopt Toulmin's (2003) model of argumentation to further refine our consideration of students' personal inferences. Toulmin's (2003) model, shown in Figure 1, includes data, a claim, and warrants. In his model of argumentation, an individual perceives data, and formulates a claim, which is a conclusion made based on data. The warrant refers to the justification that the individual uses to explain how the data support the claim. In this paper, we focus on students' personal inferences as warrant types. These personal inferences (warrants) are embedded in their arguments as they justify their own data-based claims.

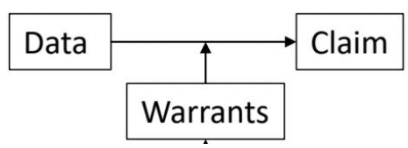


**Figure 1: Toulmin's model of argumentation (2003)**

## Research Methodology

The data for this paper are from a more extensive study (e.g., David et al., 2019, 2020; Sellers et al., 2021). In that study, we conducted individual clinical interviews (Clement, 2000) with nine students at a large public university in the USA. This paper focuses on three students, Marie, Zack, and Hannah (pseudonyms), who had completed a first-semester calculus course but had not yet taken any proof-oriented mathematics courses. Each author of this paper served as an interviewer, witness, or researcher analyzing these students' personal inferences from their verbal justification of statements in a calculus context.

The main tasks implemented during the clinical interviews and relevant to this paper are about four statements as follows:

S1: Suppose $f$ is a continuous function on the closed interval $[a, b]$, where $f(a) \neq f(b)$. Then for all real numbers $c$ in $(a, b)$, there exists a real number $N$ between $f(a)$ and $f(b)$ such that $f(c) = N$.

S2: Suppose $f$ is a continuous function on the closed interval $[a, b]$, where $f(a) \neq f(b)$. Then for all real numbers $N$ between $f(a)$ and $f(b)$, there exists a real number $c$ in $(a, b)$ such that $f(c) = N$.
S3: Suppose $f$ is a continuous function on the closed interval $[a, b]$, where $f(a) \neq f(b)$. Then there exists a real number $N$ between $f(a)$ and $f(b)$ such that for all real numbers $c$ in $(a, b)$, $f(c) = N$.
S4: Suppose $f$ is a continuous function on the closed interval $[a, b]$, where $f(a) \neq f(b)$. Then there exists a real number $c$ in $(a, b)$ such that for all real numbers $N$ between $f(a)$ and $f(b)$, $f(c) = N$.

These four statements are similar in that they are all conditional statements with the same premise, and their conclusions describe a relationship between two quantified variables, *c* and *N*. On the other hand, each of their conclusions represents different relationships between the two quantified variables. The S2 (a version of the Intermediate Value Theorem, ITV) is the only true statement among the four statements.

Without informing the students that S2 is the IVT statement, the interviewer first showed one statement at a time. Each student was then asked to interpret each statement in their own words, evaluate whether it is true or false, and explain how they could tell. Grid paper was also provided for those who wanted to sketch their example functions during the interviews. Once a student completed his/her evaluation of the four statements, the interviewer showed all four statements together and asked to compare and review their evaluations of these statements altogether.

Based on the students' utterances, gestures, and diagrams that they produced to explain their reasoning, we modeled students' arguments regarding how they would tell whether the given statements are true or false. In particular, we used Toulmin's (2003) framework to identify what students perceive (data) from the statement, their evaluation of the truth-value of the statement (claim), and how they support their claims from their data (warrants). Two types of warrants emerged from our process of analyzing students' responses to the interviewer's question about how they could determine the truth-value of the statement. One type of warrants that we found was related to students' meaning or construal of calculus concepts or their properties. Their reference to a working definition for functions and continuous functions is an example that we classified as this type of meaning-based warrants. On the other hand, the other warrants were not content-specific and were rather general rules or principles that students could apply to any content area in moving from data to claim in their arguments. In this paper, we focus on the second type of warrants, i.e., students' personal inferences that these students used as warrant types in their arguments.

## Results

While some of the students changed their evaluation of the statements, all three students determined S1-S4 as true statements in their final evaluation. Thus, we identified the *claims* in these students' arguments as identical to "the statement is true" across all four statements. On the other hand, there were differences in what constituted data of each student argument as they perceived different things from the given statements. There were also differences in their personal inferences as ways in which they justified their claims about the statements' truth. In the following subsections, we present these differences in the students' *data* and detail personal inferences that the students utilized as warrants in their arguments.

**Marie's Use of Biconditional Elimination and Generalization from a Particular**

To support her claim that S1 is true, Marie drew a graph of a function $f$ (see Fig. 2). She then explained that her function $f$ is continuous on $[a, b]$ because of no holes or jumps on the graph of $f$. She also explained $f(a) \neq f(b)$ because $f(a) < 0$ and $f(b) > 0$. She then picked a value of $c$ between $a$ and $b$ $(a < b)$ on the $x$-axis and then confirmed that its corresponding function value $f(c)$ is between $f(a)$ and $f(b)$ and thus $N$ is between $f(a)$ and $f(b)$, which seemed sufficient for her to conclude S1 is true. Referring to the same graph, Marie described her reasoning about the other three statements, S2-S4, in a similar way. For instance, explaining her reasoning about S2, she said, "So $N$ is gonna be in between here [$f(a)$] and here [$f(b)$]. And then $c$ is gonna be in between $a$ and $b$. (long pause) Okay, so $N$ is in between $f(b)$ and $f(a)$, and because it's continuous, $c$ is gonna be in between $a$ and $b$." Here, we found what she perceived from the given statements was her graph of a continuous function $f$ where $f(a) \neq f(b)$. Thus, we identified her graph as *data* in her arguments.

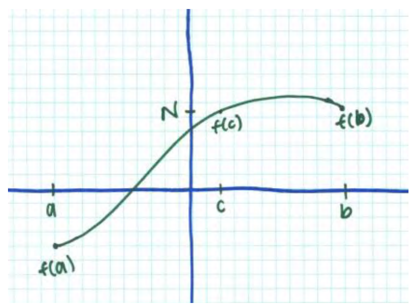


**Figure 2: Marie's graph**

Marie mainly used two inferences to support her claims. Her first personal inference is what is called *biconditional elimination*, $[(p \leftrightarrow p') \wedge p'] \rightarrow p$, in the conventional logic. Her use of this inference appeared when she replaced or substituted the premise and conclusion of the given statement to her own interpretation of them, respectively. In the case of S2, Marie interpreted the premise of the statement as a description of a function with no holes or jumps on its graph, and the conclusion as a description of the input of the function to be between $a$ and $b$ and the output of the function to be between $f(a)$ and $f(b)$. At the moment when she substituted the conclusion to these two conditions about where the input and outputs are contained, Marie neglected the quantifier words 'there exists' and 'for all.' Consequently, her interpretation of the conclusion of all four statements S1-S4 became the same. Although Marie used normatively biconditional elimination as her personal inference, her reasoning entailed invalid arguments as her evoked meaning from the given statements was not equivalent to the conclusions.

There was another inference Marie used across her arguments during the interview, which is what we refer to as a *generalization from a particular*[1], $P(x_o) \rightarrow (\forall x, P(x))$. Her function $f$ (Fig. 2) was merely an example that satisfied her evoked meaning of the given statements. However, she drew a claim that the statements are true from this particular example. Generalization from a particular is not a valid inference in conventional mathematical logic.

1 We detailed in Sellers et al. (2021) that generalization from a particular may be connected to students' quantifications.

**Zack's Use of Conjunction Elimination, Existential Instantiation, & Hypothetical Syllogism**

During Zack's interpretation of S1, the interviewer recommended that he draw a graph of a function. In response, Zack drew a graph of a linear function with a positive slope (see Fig. 3) because the graph "is just simple so I can imagine myself." However, Zack described that "$f$ will be anything, any type of a graph." In fact, he did not make his arguments based on this particular graph. His way of using a graph for illustration purposes was thus distinct from Marie's use of a graph for generalization from a particular.

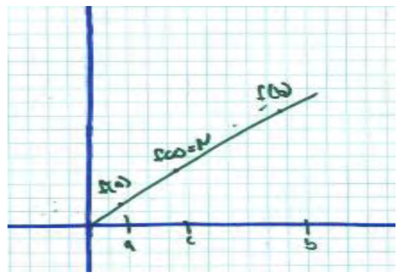


**Figure 3: Zack's graph**

Zack repeatedly pointed out a functional relationship between inputs and their corresponding outputs. Specifically, he started with *a*, *b*, and *c* as inputs and then described $f(a), f(b), f(c)$ as the corresponding outputs of a function, respectively. Zack's action-related explanation indicates that he mainly focused on the functional relationship between *c* and *N*. We thus identify his construal of such a functional relationship as *data* of his arguments. See the transcript below for his focus on the functional relationship.

> Zack on S1: When I put $f(a)$, we get that here (marks $f(a)$ on the graph) [...] $f(b)$ somewhere here (marks $f(b)$ on the graph). And then, [...] I would show it [*c*] over here (between $a$ and *b* on the *x*-axis). I would hope it [$f(c)$] would be on the same line in between the interval from $a$ to $b$ ($f(a)$ and $f(b)$ in his graph, Fig. 3). That [$f(c)$] is a real number and exists right here (between $f(a)$ and $f(b)$). Now it [$f(c)$] will be like $N$.

We noticed that Zack leveraged the functional relationship that he was envisioning (i.e., *data*) to claim that the truth value for S1 and S2 are the same. Zack attended to the quantifier words such as "for all" and "there exists" and the order differences between quantified variables $c$ and $N$ in each statement. However, his attendance to such syntactic differences between the given statements was insufficient to evoke a different construal from each statement. For instance, he explained that "this one [S2] is almost a converse statement of the previous [S1]. I think the previous [S1] was asking if $N$ existed, and this one [S2] is asking if $c$ existed. So, I will say this statement 2 is true." His wording here also supports our model of data in his arguments that he perceived the functional relationship between *c* and *N*.

Zack used several inferences to warrant his arguments to his claims that all four statements are true, and they were often compatible with normative inferences. First, he examined only the functional relationship but excluded the other conditions from the premise of the statements, such as $f$ being continuous and $f(a) \neq f(b)$[2]. We consider he subconsciously used what is called *conjunction elimination* ($p \wedge q \rightarrow p$) in conventional logic. His use of this inference

2 David et al. (2019) documented Zack's ways of thinking about the outputs of a function as spatial locations in grid paper, which explains why he excluded the condition $f(a) \neq f(b)$ from his interpretation of the premise because to him it is always the case for all functions.

differed from Marie's use of biconditional elimination: Marie replaced or substituted a condition to another as she treated them as equivalent, whereas Zack excluded some conditions from the given statement.

We also noted from Zack's explanation on S1 that he did not treat *N* as another label for an output $f(c)$, or use *N* to refer to the output value $f(c)$, until he checked that $f(c)$ is between $f(a)$ and $f(b)$. He said, "That [$f(c)$] is a real number and exists right here (between $f(a)$ and $f(b)$). Now it [$f(c)$] will be like $N$." Zack's wording here indicates that he labeled the output value $f(c)$ as $N$ as a way to justify the existence of *N*. Once Zack showed the existence of *N* from his work on S1, he used both $f(c)$ and $N$ to refer to the output value for the specific input value *c* in explaining S2-S4. We consider his labeling of *N* here as evidence of his use of what is called *existential instantiation* in conventional logic, i.e., $(x_0 \in X \ \wedge P(x_0)) \Longrightarrow (\exists x \in X(P(x))$. This inference is one of the conventional logical rules that are predominantly used in mathematics as a warrant while proving an existentially quantified statement. However, such an inference will unlikely be students' personal inference if they do not attend to the existential quantifier, as shown in Marie's case. Thus, it is noteworthy that Zack exhibited existential instantiation as his personal inference in his arguments.

Finally, we found that Zack used a chain of inferences when he used his evaluation of S1 to determine the other three statements S2-S4 to be true: He first compared the conclusion of S1 with the conclusion of S2 and insisted that if the conclusion of S1 is true, then the conclusion of S2 is true. Using his earlier justification regarding the conclusion of S1 to be true, he contended that the conclusion of S2 is also true. We consider it as an example of his use of a *logical syllogism* ($[(p \to q) \wedge (q \to r)] \to (p \to r)$) in a normative way. On the other hand, he did not provide any further warrant to support his first inference between the conclusions of S1 and S2. He just said "there is no logical reasoning behind" in his action of choosing *N* first and then choosing *c* when thinking about the conclusion of S2, and he could do the other way "choosing *c* first and then choosing *N*" as he was thinking about the conclusion of S1. To him, the conclusions of all four statements represent the same functional relationship in which $c$ is an input and $f(c)$, or $N$, is its corresponding output. Consequently, while Zack used conventional inferences normatively in his arguments, his arguments were often invalid because he construed only the functional relationship from all four statements.

**Hannah's Use of Universal Generalization, Modus Ponens, & Modus Inverse**

When the interviewer asked Hannah to evaluate S1, she drew a graph of a function as an example of a continuous function (Fig. 4, left). Hannah also drew two more graphs for discontinuous functions, which contained either a hole or a jump (Fig. 4, right). She used both graph types to explain her reasoning behind her evaluation. Although she drew these specific example graphs for continuous or discontinuous functions, she said that "[it] doesn't matter what it [the graph of a function] looks like." It is thus likely that what Hannah perceived from the given statements are generic functions satisfying or failing to satisfy the premise. We thus her generic examples as *data* in her arguments.

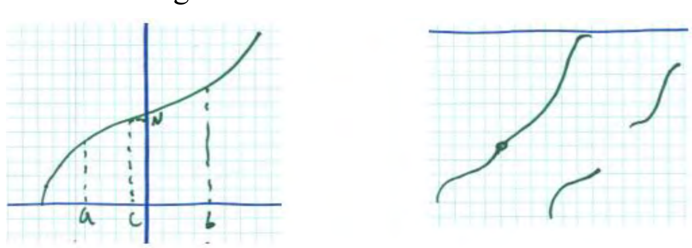


**Figure 4: Hannah's graph**

Hannah on S1: Because there is [are] no breaks, […] if you have this graph and there is a hole in it (Fig. 4, right), then that won't be continuous. So that would be an example of where this [the premise of S1] would be false. But since it's continuous, there is [are] no breaks [un]like this, or there is [are] no jumps [un]like that. (Fig. 4, left) So […] there has to be a $f(c)$ that equals to $N$, which is like, when you input any value of $x$ between $a$ and $b$, for example, $c$ (marks $c$ on Fig. 4, left), there has to be a $y$ for it, a number that $N$ equals, this roughly would be an $N$ (marks $N$ on Fig.4, left).

Hannah used several inferential warrants in normative ways to support her claims from her data. First of all, her use of continuous functions and discontinuous functions as any particular example indicates that she used what is called *universal generalization*, i.e., (for any fixed $x_0 \in X, P(x_0)) \to (\forall\, x \in X(P(x))$. We also found that Hannah used another inference that we call *modus ponens*, $(p \to q) \wedge p) \to q$, normatively in her argument for S1: She assumed the premise of S1. For her, a continuous function is equivalent to (or at least implies) that there is no break or hole in its graph. And, for her, no breaks or holes on the graph means for each input $c$ between $a$ and $b$, the output $f(c)$ is between $f(a)$ and $f(b)$.

Hannah attended to the syntactic differences between S1 and S2 and construed distinct meanings from each statement. Furthermore, she conceived that S1 and S2 would potentially represent different each-to-some relationships, i.e., S1 represents the each-$c$-to-some-$N$ relationship, whereas S2 represents the each-$N$-to-some-$c$ relationship. Her construal of different relationships from different statements characterizes her arguments distinct from Marie and Zack. However, she expressed her confusion about S3. In particular, she was not sure if the conclusion of S3 would describe a one-to-every relationship or an each-to-some relationship. Although she finally interpreted the conclusion of S3 as a description of an each-to-some relationship between $c$ and $N$, she was uncertain about her interpretation of the conclusion of S3.

*Hannah* on S3: It sounded different [from S2]. But if you think it through, it's not saying that there is only one $N$. (reading the conclusion of S2) Okay, I am confused. Is it saying there is only one $N$ or there exists a number, just any number? That's what is confusing me.

As she experienced some difficulty with S3, she focused on the sameness rather than differences between the statements, saying, "All four [statements] are equivalent because like they are just saying that there exists an input where there is a real output in this interval. So, they [S1-S4] are all saying the same thing." Here, Hannah demonstrated that some parts of the statements were distinct. However, she used her meaning of function and the input-output relationship to conclude that all four statements held the same meaning. Hannah resolved her confusion by substituting the conclusion of S3 to a conjunction of two conditions about the existence of the input and output values.

Once Hannah treated all four statements as equivalent, she also provided another way of justifying her claims about these statements, namely treating each of the given statements S1-S4 as equivalent to its inverse. She believed that she could say a statement is true by showing a statement which is what we call the inverse of a conditional statement. She thought that if a function $f$ is not continuous, then the conclusion of the given statement must be false. Therefore, if $f$ is continuous, then the conclusion of the statement is true.

Hannah: If you take a part of the statement, I guess, into two parts, it's saying it has to be a continuous function at first. Because this is discontinuous, in order to prove this statement true, you have to show that in a situation where it's not a continuous function, the second part [the conclusion] is not true. If you take a situation where like it doesn't

> meet like the basic requirements [the premise], the second part [the conclusion] can't be true for this statement to be true.

We characterize her use of such a personal inference as what we refer to in this paper as *modus inverse*, $[(p \rightarrow q) \wedge \neg p] \rightarrow \neg q$. Her use of modus inverse seemed rooted in her conception of the equivalence between a conditional statement and its inverse, i.e., $(p \rightarrow q) \Leftrightarrow (\neg p \rightarrow \neg q)$. On the other hand, the modus inverse was not aligned with normative logical inferences as there are cases in that a conditional statement is not equivalent to its inverse. Thus, her use of this inference made her assertion invalid in this case.

## Conclusion & Discussion

Our findings suggest that these calculus students do have intuitive personal inferences that are quite sophisticated and often in line with mathematical/logical convention. All three students spontaneously used various inferences rules normatively. Marie and Hannah's use of biconditional elimination, Zack's use of conjunctive elimination, existential instantiation, and hypothetical syllogism, and Hannah's use of universal generalization and modus ponens were examples of personal inferences that these students used normatively. It was noteworthy that these students utilized these logical inference rules normatively to support their claims even though they had not received any formal instruction for conventional logic or proof writing in mathematics.

While these students used several logical inferences normatively, their arguments often led to inaccurate conclusions. Partially, these students' invalid personal inferences (e.g., Marie's use of generalization from a particular and Hannah's use of modus inverse) explain how inaccurate conclusions were derived in their arguments. However, in some other cases, these students' use of content-specific warrants affected their argument to be invalid even if their personal inferences were valid. Marie's meanings for quantified variables influenced the invalidity in her arguments. Zack's location thinking (i.e., thinking of output values of a function as the spatial locations of points on the graph of a function) was also closely related to his invalid argument. We hope to investigate further how to leverage students to use their personal inferences with content-specific warrants for valid arguments in mathematics.

## References


Alcock, L. & Weber, K. (2005). Proof validation in real analysis: Inferring and checking warrants. *Journal of Mathematical Behavior*, *24*, 125-134.

Clement, J. (2000). Analysis of clinical interviews: Foundations and model viability. In R. A. Lesh & A. E. Kelly (Eds.), *Handbook of research design in mathematics and science education* (pp. 547–589). Lawrence Erlbaum.

David, E. J., Roh, K. H., & Sellers, M.E. (2019). Value-thinking and location-thinking: An empirical study of two ways students visualize points and reason about graphs in the Cartesian plane. *Journal of Mathematical Behavior*, 54, 1–18.

David, E. J., Roh, K. H., & Sellers, M.E. (2020). Teaching the representations of concepts in calculus: The case of the intermediate value theorem. *PRIMUS*, *30*(2), 191-210.

Dawkins, P. C., Roh, K. H., Eckman, D., & Cho, Y. K. (2021). Theo's reinvention of the logic of conditional statements' proofs rooted in set-based reasoning. In D. Olanoff, K. Johnson, & S. M. Spitzer (Eds.), *Proceedings of the forty-third annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp.518–527), Philadelphia, PA.

Dawkins, P., & Zazkis, D. (2021). Using moment-by-moment reading protocols to understand students' processes of reading mathematical proof. *Journal for Research in Mathematics Education, 52*(5), 510-538.

Glasersfeld, E. (1995). *Radical constructivism: A way of knowing and learning*. Falmer Press.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. Schoenfeld, J. Kaput, & E. Dubminsky (Eds.), *Research in collegiate mathematics education*, III (pp. 234-283). Providence, RI: American Mathematical Society.

Martin, G. & Harel, G. (1989). Proof frames of preservice elementary teachers. *Journal for Research in Mathematics Education*, 20, 41-51.

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

National Governors Association Center for Best Practices, & Council of Chief State School Officers (2010). *Common core state standards initiative: Common core state standards for mathematics*. Washington, DC: Authors.

Sellers, M. E., Roh, K. H., & Parr, E. D. (2021). Student quantifications as meanings for quantified variables in complex mathematical statements. *Journal of Mathematical Behavior*, *61*, 1–16.

Shipman, B., & Shipman, P. (2013). Simple yet hidden counterexamples in undergraduate real analysis. *PRIMUS*, 23, 111-120.

Stylianides, A., Stylianides, G., & Phillipou, G. (2004). Undergraduate students' understanding of the contraposition equivalence rule in symbolic and verbal contexts. *Educational Studies in Mathematics*, *55*, 133-162.

Toulmin, S. (2003). *The uses of argument* (Updated Ed.). Cambridge University Press.