Journal of Mathematical Behavior 59 (2020) 100794

Contents lists available at ScienceDirect

Journal of Mathematical Behavior

journal homepage: www.elsevier.com/locate/jmathb

Eliminating counterexamples: An intervention for improving adolescents' contrapositive reasoning

David A. Yopp

Departments of Mathematics and Curriculum and Instruction, University of Idaho, 875 Perimeter Drive MS 1103, Moscow, ID 83844-1103, United States

ARTICLE INFO

Keywords:
Reasoning
Proving
Contrapositive proving
Teaching intervention
Viable argument

ABSTRACT

Students' difficulties with contrapositive reasoning are well documented. Lack of intuition about contrapositive reasoning and lack of a meta-argument for the logical equivalence between a conditional claim and its contrapositive may contribute to students' struggles. This case study investigated the effectiveness of the *eliminating counterexamples* intervention in improving students' ability to construct, critique, and validate contrapositive arguments in a U.S. eighth-grade mathematics classroom. The intervention involved constructing descriptions of all possible counterexamples to a conditional claim and its contrapositive, comparing the two descriptions, noting that the descriptions are the same barring the order of phrases, and finding a counterexample to show the claim is false or viably arguing that no counterexample exists.

1. Introduction

Indirect reasoning about mathematical generalizations arises spontaneously in mathematics courses for students of all ages (Antonini, 2003, Antonini & Mariotti, 2008; 2004; Freudenthal, 1973; O'Brien, 1972; Reid and Dobbin, 1998). *Indirect reasoning* can involve formal or informal reasoning akin to proof by contrapositive or proof by contradiction. In less formal terms, *indirect reasoning* can be found in proving activities where arguers consider what occurs when the conclusion of a conditional claim[1] is not satisfied. *Contrapositive reasoning* occurs when arguers attempt to show that all cases of *not the conclusion* of a conditional claim imply the conditions are not satisfied or when arguers justify *not-the-conclusion-implies-not-the-conditions* arguments as using a viable method/mode of proof/viable argumentation.

The current article reports findings from an intervention aimed at improving eighth-grade students' contrapositive reasoning. In this article, *eighth grade* refers to the U.S. system, students around age 13. Contrapositive reasoning can be important in eighth-grade mathematics because this reasoning can be used to learn about relationships between two mathematical properties by considering what occurs when one property is not satisfied. In the United States, where this study occurred, *Common Core State Standards for Mathematics* (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010), the recommended standards in most U.S. states at the time this intervention, do not mention contrapositive reasoning. However, the standards state that U.S. eighth-grade students will "use informal arguments to establish facts about … the angles created when parallel lines are cut by a transversal" (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010, p. 56) and will explain a proof of the converse of the Pythagorean theorem. Both these content areas can involve contrapositive reasoning.

E-mail address: dyopp@uidaho.edu.

[1] Throughout this article, *conditional claim* is used to refer to a mathematical generalization that is either implicitly or explicitly quantified.

https://doi.org/10.1016/j.jmathb.2020.100794
Received 30 September 2018; Received in revised form 7 November 2019; Accepted 7 June 2020
Available online 26 June 2020
0732-3123/

For example, a student addressing the claim that *when parallel lines are cut by a transversal, interior angles on the same side of the transversal sum to 180 degrees* can be asked to consider what happens when two lines are cut by a transversal but the measures of a pair of interior angles on the same side of a transversal sum to less than 180 degrees. The student might acknowledge that these properties make the lines "angle in" toward each other, causing the lines to eventually intersect. Perhaps the student acknowledges that a pair of angles with measures summing to less than 180 degrees leaves "room" for another third angle to create a triangle with a vertex where the lines intersect. This informal observation can be summarized as illustrating that *if a pair of interior angles on the same side of a transversal sum to less than 180 degrees, then the lines are not parallel.* Using contraposition reasoning, this observation can be leveraged to conclude that the original claim is true. However, to draw such conclusions *and* affirm the overarching method/mode of argumentation as viable, the student likely needs training in the relationship between a conditional claim and its contrapositive.

In fact, previous studies seem to suggest that the last sentence is true. Studies have demonstrated that justifications for the relationship between a conditional claim and its contrapositive are often not in place and that students tend to do poorly on indirect reasoning tasks. In previous studies, students failed to recognize indirect arguments as proofs (Antonini, 2004) and struggled negating conditions and conclusions (Leron, 1985). Students familiar with indirect proving methods have even rejected indirect proofs as operating in absurd worlds (Antonini & Mariotti, 2008).

Interventions for improving students' proof, proving, and viable argumentation practices and knowledge are sparse and are even rarer for contrapositive proving and reasoning (Stylianides, Stylianides, & Weber, 2017). This may be due in part to the difficulty in designing proof, proving, and viable argumentation interventions, particularly for eighth-grade students. In fact, envisioning indirect reasoning approaches and justifications that are both accessible to eighth-grade students and consistent with what students will learn later can be challenging. Eighth-grade students may not have access to classical logic, where logical equivalences such as those between conditional claims and their contrapositives are established. Interventions based in the approaches found in undergraduate introduction-to-proof textbooks (where contrapositive proving is often taught) may not be helpful because many of these textbooks develop the logical equivalences between a conditional claim and its contrapositive using truth tables and propositional logic (Dumas & McCarthy, 2007; Esty & Esty, 2013; Krantz, 2002) and extend the results to predicate logic by viewing propositional statements as formulas in a variable x predicated by quantifiers (e.g., Dumas & McCarthy, 2007). These formal approaches are not consistent with what is currently taught in U.S. eighth-grade classrooms and may be ineffective anyway (Antonini, 2004).

Moreover, the self-reported practices and learning experiences of expert mathematicians also may be unhelpful for developing interventions. Literature describing how mathematicians justify contrapositive argumentation is sparse, and it is difficult to determine how experts came to know about contrapositive reasoning. Experts trained in classical logic may find the link between $p \rightarrow q$ and *not* $q \rightarrow$ *not* p so obvious it requires no explanation (Antonini, 2004).

Fortunately, there may be other, mathematically less formal ways of approaching this problem. Antonini (2004) posits that students might benefit from producing informal arguments of the form "If [blank] were not so, then [blank] would happen," then guiding students toward the contrapositive structure, and existing psychology literature describes informal indirect reasoning of people not necessarily trained in formal logic. The premise of the current article is that people's informal indirect reasoning as described by prior researchers can be leveraged to develop and justify indirect arguments and distinguish these arguments from direct arguments.

1.1. Introduction to the eliminating counterexamples approach to contrapositive reasoning and its theoretical origins

The intervention described in this article was motivated in part by considering ways in which peoples' everyday reasoning abilities can be leveraged toward reasoning that better aligns with mathematical notions of proof, proving, and viable argumentation. One scheme from the psychology literature that seeks to describe people's everyday reasoning is the *pragmatic reasoning scheme* (Cheng & Holyoak, 1985). This scheme suggests that people reason using context-referenced reasoning rules and the natural-language meanings of words like "only if." Another scheme is the *mental models reasoning scheme* (Johnson-Laird, 1983), which suggests that people build alternative models to a conditional claim and then leverage these models to eliminate the possibility of counterexamples. These psychological models attempt to explain reasoning, but I assert that these models can be reenvisioned to offer ideas about how to improve contrapositive reasoning.

The *eliminating counterexamples (ECE)* approach (Yopp, 2017) to contrapositive reasoning sought to do this by basing the approach, in part, on a reenvisioning of one of the mental-models reasoning scheme's tenets—that reasoners may construct alternative models to a conditional claim's conclusion when reasoning about the truth of a claim. The ECE intervention involved a related approach: developing a description of all possible cases of counterexamples to a conditional claim (i.e., a hypothetical model for alternatives to a conditional claim) and then, when the claim is true, demonstrating that no example satisfies this description. The ECE approach also leveraged tenets from the pragmatic scheme. Arguers can make sense of the approach both in and out of context through the meaning of the phrase "only if"—a conditional claim is true only if no counterexample exists; lines are parallel only if alternate interior angles sum to 180 degrees.

While contrapositive reasoning involves a reasoning rule that when every object satisfying not-the-conclusion also has the property not-the-conditions, the original claim is true, a justification for why this mode of reasoning supports the original claim is needed. This justification can be constructed from the ECE approach by acknowledging that conditional claims are true if and only if no counterexamples exist. Arguers can note that showing not-the-conclusion implies not-the-conditions demonstrates that cases of the conditions and not-the-conclusion cannot exist. Alternatively, the arguer can note that the collection of counterexamples to the original conditions claim is equal to the collection of counterexamples to its contrapositive, which informally supports the logical equivalence of these two forms. Eliminate counterexamples to either one of these forms, and you have eliminated counterexamples to the other. The pragmatic reasoning scheme can come into play when the arguer uses phrase like “only if” when describing reasoning about the existence of counterexamples. When considering possible counterexamples to p implies q — p and not q —an arguer may reason that not q occurs only if p does not happen.

Yopp (2017) argued that the ECE approach is consistent with learning from propositional and predicate logic, and yet the approach is distinct from how contrapositive proof and proving are presented in undergraduate mathematics textbooks. Yopp reported that this approach was both accessible and useful to a particular mathematically advanced eighth-grade student as she constructed, critiqued, and validated contrapositive arguments.

1.2. The research question addressed

In this study, I extended the findings of Yopp (2017) to a larger and more academically diverse group of students enrolled in eighth-grade mathematics. I explored the changes that occurred among students' responses to assessment items that involved either direct or indirect reasoning after the students experienced the ECE intervention, an intervention that teaches the ECE approach. The following research question was addressed: In what ways do a group of students enrolled in a U.S. eighth-grade mathematics course change their contrapositive reasoning in response to the ECE intervention?

2. Literature on contrapositive proving and argumentation

Students' difficulties/discomfort with indirect proof and proving are well established (Antonini & Mariotti, 2008; Harel & Sowder, 1998; Leron, 1985). Yet Brown (2018) found that students' views about whether an indirect proof is more convincing than a direct argument for the same generalization were more nuanced and subjective than indicated previously. Students' views were dependent on the argument's complexity and student's familiarity with the context. Intuition may also play a role in whether students accept indirect arguments as viable/proof. Students may express accurate intuitions about alternative forms of a conditional claim but not acknowledge logical equivalence between a conditional claim and its contrapositive (Antonini, 2004).

Little is known about ways to improve students' contrapositive proof and proving knowledge and practices, but Antonini and Mariotti (2008) posited that students may need a meta-theorem to understand contrapositive reasoning's validity. Loosely defined, meta-level reasoning is reasoning about reasoning. To Antonini and Mariotti, a meta-theorem for contrapositive proving consists of (1) a meta-statement for the equivalence between a conditional claim and its contrapositive, (2) a meta-proof of the meta-statement, and (3) a meta-theory that serves as the theoretical basis for the meta-proof. Building upon this suggestion, Yopp (2017) replaced Antonini and Mariotti's mathematical formal notions of a meta-theorem for contrapositive proving with the less formal justifications found in the ECE approach. Yopp (2017) found that these meta-level justifications and the ECE approach were useful to a particular high-achieving eighth-grade mathematics student as she constructed and critiqued contrapositive arguments.

3. Literature on classroom-based interventions in the area of proof, proving, and argumentation

The phrase *classroom-based interventions in the area of proof* has been used to describe interventions designed to improve understanding or use of proof in formal mathematics learning settings at any grade level (Stylianides et al., 2017). According to Stylianides et al. (2017), “The number of such studies is small and acutely disproportionate to the number of studies that have documented problems of classroom practice [in the area of proof] for which solutions are sorely needed” (p. 253).

A few classroom-based intervention studies demonstrated that students' perspectives on what constitutes proof of a generalization can be shifted from acceptance of empirical or intuitive justifications to desires for more secure justifications (Mariotti, 2013; Stylianides & Stylianides, 2009). A few other classroom-based interventions demonstrate the possibility of improving students' understandings of norms for *proof* (e.g., Alibert & Thomas, 2002; Goos, 2004).

Instruction and teacher input on proof and proving practice or strategies appear to be important in improving students' proof and proving responses. Komatsu (2017) found that the proving tasks with diagrams plus teachers' actions encouraged students to discover cases that reject proofs, refute statements, modify proofs, disclose hidden conditions, or restrict domains of the statements. Yopp (2015) demonstrated that explicit instruction on (1) the mathematical register for stating existence claims and generalizations, (2) frameworks/standards for generating types of arguments for each type of claim, and (3) metamathematical practices like those in Lakatos (1976), such as searching for counterexamples and restricting claims to avoid counterexamples, can improve preservice elementary teachers' responses to false generalization. Perhaps most relevant to the current study, Jahnke and Wambach (2013) reported on an intervention in Germany that improved eighth-grade students' understanding that proofs depend on assumptions. Despite this progress, middle-grade interventions in the area of proof are limited, and not all intervention studies have demonstrated positive results. Fan, Qi, Liu, Wang, and Lin (2017) reported on an intervention involving Chinese eighth-grade students who received instruction on drawing auxiliary lines to solve proving tasks. Fan et al. found no effect overall but argued that the treatment group did better than the comparison group on “high-level cognitive” tasks.

4. Theoretical framework

As mentioned above, Yopp (2017) explored an intervention for improving a particular student's contrapositive reasoning using the ECE approach, which takes a less formal approach than developing a meta-theorem as suggested by Antonini and Mariotti (2008). The approach involved working with contrapositive reasoning at two levels: the level of a particular generalization (arguing that the particular generalization is true) and the metamathematical level (justifying contrapositive argumentation as a viable mode of argumentation).

4.1. More on the theory behind the eliminating counterexamples intervention

At the level of a particular generalization, the ECE approach involves developing a description of the collection of all possible counterexamples to a generalization (in conditional claim form) and arguing that no example satisfies the description. At the metamathematical level, the approach involves at least one of two activities, (1) developing descriptions of the possible counterexamples to a conditional claim and its contrapositive and noting the two descriptions are the same, barring order of phrases, and generalizing the findings, or (2) noting that contrapositive argumentation eliminates the possibility of counterexamples by demonstrating that the properties of the conditions and those of not-the-conclusion cannot occur simultaneously. The later may occur in less than mathematically formal language, using phrases such as “only if,” as described in the pragmatic reasoning scheme literature (Cheng & Holyoak, 1985). Either activity leads to a notion of contrapositive argumentation as *eliminating the possibility of counterexamples* to the original generalization (conditional claim).

The ECE approach and corresponding intervention were developed with reflection on nearly a decade of research on proof, proving, and viable argumentation in mathematics learning environments)Yopp, 2011, 2014, 2015, 2017). Critical to the intervention’s development was reflection on potential roles of proof/proving in mathematics instruction. Yopp (2011) explored how some mathematics professors use proof in their classrooms. This study enabled a richer understanding of the possible uses of proof and proving in mathematics teaching. Most compelling were uses for increasing understanding of theorems being proved and of the definitions of the mathematical objects in those theorems. Yopp (2011) also developed a distinction between proving to establish horizontal connections and proving to establish vertical connections:

Systematization or axiomatization builds an expanding theory based on axioms or definitions with intent to order results, which can be described as a “vertical” structure. Building connections among mathematical ideas can be “horizontal” in that proof can connect two seemingly unrelated mathematical results whose proofs do not directly depend on each other. A person connecting mathematical ideas may have no interest in developing a coherent mathematical theory and, instead, may only be interested in the mathematics ideas at hand. (p. 123)

Learning environments not focused on systemization can take advantage of opportunities to connect and coordinate ideas without being precise about which result was established first. This acknowledgment was important to the ECE intervention’s design because current U.S. eighth-grade standards, at the time of this study, do not call for systemization (NGACBP & CCSSO, 2010) but do ask that students make arguments and learn proofs. Consequently, the ECE intervention places emphasis on acknowledging that arguments appeal to already-learned results, not axiomatic systems development.

The ECE intervention also employed Yopp's (2015) idea that students need explicit instruction on the language of proof, the technical skills needed to construct proofs, and the canonical practices and beliefs of mathematicians. Yopp (2015) found that students benefited from explicit instruction on the mathematical register (e.g., *for all, if-then, there exists*), on closed versus open sentences, on counterexamples and their meanings, on standards for proving existence claims and generalizations, on distinctions between expressing one’s ideas/thoughts and developing viable arguments or proofs, and on the practice of exception barring (Lakatos, 1976). Yopp, 2015Yopp (2015) found that students also benefited from developing descriptions of the collection of all possible counterexamples to a generalization when participating in metamathematical practices similar to those in Lakatos (1976).

The ECE intervention encouraged example uses for various purposes, including understanding what a claim says and providing a referent for proving the claim. This encouragement was found in Yopp (2014) finding that students in an undergraduate introduction-to-mathematical-proofs class used examples in productive ways when exploring conjectures and knew that their examples did not prove the generalization they explored. Students used examples to develop shared meanings of a task, communicate meanings/approaches to other students, express/resolve disagreements about interpretations/approaches, construct existence arguments, and navigate a generalization with a nested existence quantifier.

The ECE intervention also embraced Yopp and Ely's (2016) assertion that arguments leveraging an example as a referent (e.g., generic examples) can qualify as formal proof. Yopp and Ely (2016) positioned arguments that leverage examples on the same level of rigor/sophistication as arguments that rely on variable-based expression, equations, or diagrams, provided that the example-based arguments satisfy a collection of criteria that ensure the argument does not appeal to any feature of the referent not shared by all cases in the domain of the generalization being proved. Authors placed less emphasis on what example/referent is used in an argument than on how the example is used.

Finally, the intervention acknowledged that students may acquire sophisticated ways of thinking about proof, proving, and viable argumentation that may not occur to experts. As in Yopp (2017), the ECE intervention assumed students’ emerging notions of proof, proving, and viable arguments can be leveraged into a framework accessible and useful to learners yet distinct from how experts might present these notions.

4.2. Proof, proving, and viable argumentation

A debate about *proof* and related terms in mathematics education literature is beyond this article’s scope. I offer instead a summary of how I use these terms.

For this study, *critiquing* an argument refers to assessing its viability, including but not limited to identifying its strengths and weakness, methods/modes of argumentation, logical steps, referent use, and flaws. *Validating* an argument refers to sanctioning or affirming part or all of an argument as viable or a proof, or a step in an argument as valid or sound. *Justifying* refers to supporting or defending the viability of any of the argument's features.

Argument refers to either the support for a claim or both the claim and its support. The term contains no evaluation as to whether the argument can be taken as proof. *Argumentation* refers to any means used to support or construct a claim. *Viable argument* refers to an argument that can be taken as proof but may have features that are less than formal, such as implicit inferences or intuitive argumentation/proof approaches that may not align with or attend to canonical methods or modes of proof. For conditional claims, this means that the argument attends to all cases in the domain of the claim and that the argument demonstrates, perhaps using informal language, reasoning, and representations, that there are no examples that satisfy both the conditions and not the conclusion.

5. Methods

5.1. Interventions that involve teaching

The intervention described in this study involved teaching, and the design adapted practices from an approach that Simon (1995) calls a *teaching experiment*. The field has not drawn a clear distinction between the terms *intervention* and *teaching experiment*, with teaching experiment being a somewhat loaded term associated with a particular methodology. Teaching experiments create environments where students' learning is explored and hypothetical models for teaching and learning are elaborated (Simon, 1995). In Simon (1995), instruction was modified and adapted to meet students' needs as the students' responses were analyzed and the hypothetical model was updated. Interventions often make explicit the types of instruction students will experience, the intended changes anticipated in students, and the mechanisms used to produce the change. Interventions might seek to compare students' learning to a predetermined collection of anticipated outcomes. Teaching experiments might do the same but often seek to document students' activities and conceptions as they engage in materials, activities, and discussions. Despite the distinctions, it is possible to find situations where the two methodologies coincide.

Stylianides and Stylianides' study (2009) stands out as coordinating the two methodologies effectively. This particular study made clear the activities and instruction students experienced, the intended outcomes, and the mechanisms anticipated to cause the change in students' knowledge and actions. Tasks were designed to create cognitive conflict among preservice elementary teachers' notions of proof. Students initially accepted empirical support for generalizations as "proof," only to find one or more counterexamples among the examples not tested. The intervention included tasks and delivery mechanisms designed for the particular change, and data were analyzed against this anticipated change. The current study adopts a similar approach.

5.2. Sample

The study involved 15 students from an eighth-grade mathematics class in a public charter school in the U.S. Northwest. Charter schools are publicly funded but managed by private organizations. U.S. eighth-grade students are typically around age 13, but younger students with high mathematics achievement are often placed in higher-grade mathematics classes without changing the students' overall grade-level status.

In this study, two participating students were seventh graders (age 12), two were sixth graders (age 11), and the remainder were eighth graders (age 13). These students represented diverse levels of prior mathematics achievement, as measured by annual Smarter Balance Assessment Consortium grade-level assessments (http://www.smarterbalanced.org/). See Table 1.

Table 1
Characteristics of the Students in This Study.

Student	Grade level	Previous grade-level achievement	Current grade-level achievement
Student A	8	Advanced	Advanced
Student B	8	Basic	Basic
Student C	8	Not available	Basic
Student D	8	Proficient	Basic
Student E	8	Basic	Basic
Student F	8	Proficient	Proficient
Student G	8	Not available	Proficient
Student H	8	Not available	Below basic
Student I	7	Advanced	Advanced
Student J	6	Advanced	Advanced
Student K	6	Advanced	Advanced
Student L	7	Advanced	Advanced
Student M	8	Proficient	Proficient
Student N	8	Proficient	Proficient
Student O	8	Below basic	Basic

5.3. [Project 1, name blinded]

Before the ECE intervention, students experienced significant training in viable argumentation through the [Project 1] intervention. This larger intervention was designed to make viable argumentation a practice for learning eighth-grade mathematics content and was a significant part of the students' regular mathematics classroom instruction throughout the school year (late August–late May).

The [Project 1] intervention was predicated on the notion that high-quality mathematics instruction supports viable argumentation as a daily feature of teaching and learning and a regular feature of assessment. A typical lesson introduced students to, or reminded students of, a precise mathematical definition related to the content taught. Definitions were quantified and written in if-and-only-if form. Students were asked to develop viable argument in response to prompts by writing general and existence arguments with explicit reference to definitions of the objects involved and other prior results.

The students' regular classroom teacher was trained in and agreed to adopt a mindset of *teaching and learning with and through viable argumentation* in which mathematics concepts were taught and learned by participating in viable argument activities, including developing viable arguments for or against mathematical claims and, at times, developing claims. The teacher's training included two years of coaching by the author and a two week professional development lead by the author. The eighth-grade lessons the teacher implemented involved making/addressing generalizations and existence claims. Teachers and students were to be explicit about the domains of their claims and make the domains of their generalizations as general as their data or conceptual insights supported. Students and teachers were expected to (a) use mathematical language (e.g., for all, if-then, there exists, etc.) to make existence claims and generalizations and to make note of whether generalizations have finite or infinite domains, (b) be skeptical of empirical evidence (Brown, 2014; Stylianides & Stylianides, 2009), (c) be conscious about whether they were generalizing based on patterns in their results or patterns in their processes (Harel, 2001), and (d) search for conceptual insights (Yopp, 2015); Sandefur, Mason, Stylianides, & Watson, 2013) that structurally linked the conditions to the conclusion and explained why the claim holds for all cases.

During the [Project 1] intervention, students practiced direct proving, showing every case of the conditions had the conclusion by creating logical chains, each implying logical necessities; contrapositive proving; contradiction proving; proving by exhausting all cases; constructive existence proving; and pragmatic and model-based reasoning resulting in less-than-formal arguments. Eliminating counterexamples was emphasized as a goal of arguments for general claims but not as extensively as in the ECE intervention that followed.

5.4. The eliminating counterexamples intervention

The ECE intervention was designed to extend the [Project 1] intervention by giving students intensive experiences in a collection of practices associated with contrapositive proving/argumentation. The conceptions the intervention attempted to build were based in the intermediate conceptions (ICs) developed in Yopp (2017) documentation of a 13-year-old student's learning of the ECE approach. Prerequisite to developing the ICs was acceptance of two definitions related to proof: A generalization is true if and only if no counterexamples exist; a generalization is false if and only there exists a counterexample.

IC0. Descriptions of possible counterexamples can be used to develop arguments for or against counterexamples;

IC1. Collections of all possible counterexamples can be described by their properties, even when counterexamples are impossible;

IC2. The collection of counterexamples for a conditional claim and the collection of counterexamples for the claim's contrapositive are equivalent;

IC3. Arguments showing *not the conclusion implies the conditions are impossible* argue directly for the contrapositive and indirectly for the original claim.

The ECE intervention included direct instruction on the ICs and activities that encouraged practices associated with the ICs. The ECE intervention lasted 17 days during May and included 17 activities/tasks (see Table 2 for examples). One practice associated with ICs 1–2 was to describe, using words, symbols, or diagrams, the collection of all possible counterexamples to a generalization believed

Table 2
Sample of Activities/Tasks (and Their Purposes) From the Eliminating Counterexamples Intervention.

Sally claims that the function $f(n) = 2^{(2^n)} + 1$ outputs a prime for all natural number inputs. Develop a viable argument for or against Sally's claim.*
Purpose: Revisit the notion of counterexample and skepticism about generalizations. Introduce IC0 for false claims.
Consider the claims: If a natural number n is divisible by 3, then n is divisible by 6. If a natural number n is divisible by 6, then n is divisible by 3. (a) Give three examples that meet the conditions of each claim. (b) Give three examples that do not meet the conditions of each claim. (c) Write (in words) a general description of all possible counterexamples to each claim. (d) Viably argue for or against the possibility of counterexamples.
Purpose: Provide students with experiences describing counterexamples and experiences arguing for or against the existence of counterexamples.
Viably argue for or against the claim: If $x > 0$ and $x^2 > y^2$, then $x > y$. [Scaffolding present. Students prompted to describe possible counterexamples to the claim and to describe possible counterexamples to the contrapositive of the claim. Students prompted to compare these descriptions and then prompted to determine whether a counterexample to either claim exists.]
Purpose: Reinforce IC0–IC2. Introduce IC3 and IC4.
Determine if the claim is true or false and provide a viable argument for your response: If $x + y > 10$, then $x > 6$ or $y > 4$. [No scaffolding prompts provided.]
Purpose: Reinforce IC0–IC4.

Notes. IC = intermediate conception.
* $f(5) = 4{,}294{,}967{,}297 = 641 \cdot 6{,}700{,}417$ provides a counterexample to Sally's claim.

Table 3
Pre/Post-Eliminating Counterexamples Assessment.

Item number	Item*	Item type
Item 1	Develop a viable argument for or against the claim: If $ab = 0$, then $a = 0$ or $b = 0$. Is your argument viable? Explain why or why not.	Argument construction
Item 2	Evaluate the argument: **Claim:** If a and b are odd counting numbers, then $a + b$ is even. **Support:** Let a and b be odd counting numbers. Then a can be written as $2k + 1$ and b can be written as $2n + 1$, where k and n are counting numbers. The sum $a + b = 2k + 1 + 2n + 1 = 2k + 2n + 2 = 2(k + n + 1)$ shows that $a + b$ is even. (a) Is the argument viable? Explain why or why not. (b) What method of argumentation is the arguer using? Explain. (c) What would the arguer need to do to make the argument more viable?	Critique direct argument
Item 3	Evaluate the argument: **Claim:** Let n be a counting number. If n^2 is even, then n is even. **Support:** Let n be an odd counting number. Then n can be written as $2k + 1$, where k is an integer. Squaring, $n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + k) + 1$. (a) Is the argument viable? Explain why or why not. (b) What method of argumentation is the arguer using? Explain. (c) What would the arguer need to do to make the argument more viable?	Critique indirect argument
Item 4	Evaluate the argument: **Claim:** If $x < 0$, then $x^2 > 0$. **Support:** If $x < 0$, then $-x>0$. In other words, $-x$ is a positive number. Because the product of two positive numbers is also positive, $(-x)(-x) > 0$. Thus, $(-x)(-x) = x^2 > 0$. (a) Is the argument viable? Explain why or why not. (b) What method of argumentation is the arguer using? Explain. (c) What would the arguer need to do to make the argument more viable?	Critique direct argument
Item 5	Evaluate the argument: **Claim:** Let x be any real number. If $x^2 \neq y^2$, then $x \neq y$. **Support**: Let $x = y$. Then $x^2 = y^2$. (a) Is the argument viable? Explain why or why not. (b) What method of argumentation is the arguer using? Explain. (c) What would the arguer need to do to make the argument more viable?	Critique indirect argument

* The domains of some variables in the claim or support are implicit. For example, the convention is to take a and b in Item 1 as placeholders for any real number (or generic representations of all real numbers) and to take k and n in Item 2 to be natural (counting) numbers. Additional, this allowed opportunities to assess students' interpretations of these variables and the numbers represented.

to be true. The activities progressed from introducing one or more of the ICs, starting with IC1 and progressing in order until all four ICs were presented at the top of the page of the activities. Activities early in the intervention included scaffolding prompts that encourage practices associated with the ICs (e.g., describe all possible counterexamples). As time passed, scaffolding prompts were removed until students were given only a generalization and the prompt "viably argue for or against the claim."

5.5. Data collection

Students' knowledge and understanding of direct/indirect argumentation, the distinction between these two argument types, and meta-level justifications for their argument approaches were assessed using the same pre/post-ECE assessment (Table 3). The same assessment was used pre/post because valid and reliable parallel assessments were not available.

Classroom episodes during the ECE intervention were videotaped. Student written work during the episodes was collected.

5.6. Data analysis

Student assessments were scored/categorized using the *critiquing arguments* and *constructing and critiquing arguments* rubrics, described below. The rubrics were constructed by the author specifically for this study. The rubrics were developed based on the features of the ECE approach and intervention and the data generated from this study. The rubrics sorted responses into four levels of viable argumentation, from Level 3, *viable critique/viable argument and critique* (highest), to Level 0, *little or no evidence of viable critique/viable argumentation* (lowest). The *constructing and critiquing arguments* rubric categorized responses to Item 1 on the pre/post-assessment; the *critiquing arguments* rubric categorized responses to Items 2–5.

A Level 3 response on the *critiquing arguments* rubric, *viable critique,* included correct identification of the argument's method/mode, justification for the method's validity, and discussion of the argument's context-level details. The latter means that the student's response attended to the particular mathematical objects, their structures, and their meanings, mentioned in the claim.

A high-level response to Item 3 on the pre/post-assessment might note that the argument presented is a contrapositive argument (the method/mode) because it shows that every case of *not the conclusion* also has the properties of *not the conditions.* To justify the method's validity, a response might note that an argument of this type makes counterexamples, cases of the conditions and not the conclusion, impossible. A Level 3 response must also note some context-level details that make the argument viable. These can include, but are not limited to, noting that the algebraic manipulations or variable representations presented in the argument show that all cases of the conditions (or not the conclusion) have the properties described in the conclusion (or not the conditions).

A Level 2 response, *emerging viable critique*, included many features of a Level 3 response but was lacking in at least one of the three areas mentioned above. A Level 1 response, *elements of a viable critique*, included at least one feature of a Level 3 response but was lacking in the other two. A Level 0 response included little or no evidence of a viable critique.

The *constructing and critiquing argument* rubric was developed specific to Item 1 of the pre/post-assessment. A Level 3 response constructed a viable argument and critiqued the argument constructed appropriately.

For example, an arguer might describe the possible counterexamples to the claim as numbers a and b that have the properties $ab = 0$ and $a \neq 0$ and $b \neq 0$. The arguer might then note that examples satisfying this description are impossible because two nonzero numbers multiplied together result in a number that is also nonzero. This arguer might note the argument is viable because it shows it is impossible to have both properties in the description of possible counterexamples simultaneously.

A Level 2 response constructed a nearly viable argument and critiqued the argument appropriately but lacked sufficient details to score at Level 3. A Level 1 response constructed an argument with a viable idea or presented a viable conceptual insight that could be developed into a viable or nearly viable argument. A Level 0 response expressed no evidence of viable argumentation.

The rubrics were develop through an iterative process that began with theoretical rubrics that were revised as the data was analyzed and sorted. Theoretical rubrics were first drafted with anticipated student responses. These rubrics were researcher/teacher driven, based on argumentation approaches and justification approaches taught to students in the [Project 1] and the ECE intervention. Student responses to Items 2–5 were first organized into four groups depending on the degree to which the responses aligned with those anticipated. During this stage, responses that used language taught during the interventions (e.g., direct/indirect argument, contrapositive argument, eliminating counterexamples) were likely to be placed in higher-level categories.

As the literature review in this article illustrates, prior research on eighth-grade mathematics students' ability to communicate their reasoning suggests that students are unlikely to communicate their reasoning as well as mathematicians. While initial data analysis involved looking for responses that conformed to those theorized and taught in this study, such as using terms like "contrapositive" or "eliminate counterexamples" or applying a method/mode of argumentation presented in the teaching, a next stage of data analysis used meaning analysis, where natural language usage was valued even if language was not canonical or not anticipated in the theoretical framing of responses.

Next, the responses were sorted into groups that identified the argument as viable or not viable. Responses stating that indirect arguments (Items 3 and 5) were not viable because the arguments did not address the objects in the claim were grouped together.

Then, responses presenting justifications were sorted. Emphasis was placed on whether the response mentioned the viability of the method/mode of argumentation used in the argument and whether the response included comments about the viability and soundness of the argument's context-level features.

Next, responses that used approaches not explicitly or implicitly taught in [Project 1]/ECE interventions were identified. These approaches included context-based pragmatic reasoning, model-based reasoning, and other methods not explicitly taught. These responses were organized into groups based on the quality of elaborations and justifications of the modes of argument reported by the students.

Throughout this process, the rubrics were continuously updated until the rubrics became stable, meaning as responses were reevaluated, no new additions to the rubrics emerged. The responses were then rescored using the rubrics until scores became stable, meaning the score was the same as the one given in the previous iteration.

6. Findings

6.1. Constructing arguments

Only Item 1 asked students to construct a viable argument. In general, pre/post-assessment scores improved and the content of the responses changed pre/post. Six of fifteen students improved their scores pre/post-assessments in response to Item 1 (Table 4). The greatest categorical improvements on Item 1 resulted from changes in students' reasoning approaches and changes in students' discussions of their approaches. On the pre-assessment, nine of 15 students presented converse errors, all scoring at Level 0. On the post-assessment, only five students presented converse errors, all scoring at Level 0.

Student M's pre-assessment response was representative of those making converse errors:

Develop a viable argument: "f [foundation]: $x \bullet 0 = 0 \bullet x = 0$. N.L. [narrative link]: Anything multiplied by 0 is equivalent to 0." Argument viable? "Anything multiplied by 0 is equivalent to 0." (Student M, preassessment response, Level 0, converse error.)

In fact, only two of the 15 students gave responses to Item 1 on preassessment that contained reasoning that could be leveraged toward a viable argument. One of these two students, Student L, constructed the following response:

Develop a viable argument: "Claim: If $ab = 0$, then $a = 0$ or $b = 0$. Foundation: $b = 0$; $a \times 0 = 0$. $a = 0$; $0 \times b = 0$. Narrative Link: $ab = 0$ can be simplified into $a \bullet b = 0$. If a or b are positive numbers, then the only number it could be multiplied by to get 0 is 0." Argument viable? "No the argument is not viable. This is because I only justify that a or b must equal 0 if they are both positive numbers. Since my claim's domain applies to all numbers but I only prove the claim for positive numbers, then the argument is not viable." (Student L, preassessment response to Item 1, Level 2. pragmatic reasoning and perhaps reasoning with models)

Student L's response was restricted to positive numbers, but the discussion of what must occur if one of the variables in the conclusion is not zero provided a path toward viable argumentation. From the height of a trained mathematician's perspective,

Table 4
Item 1 Categorical Data, Pre/Post-Eliminating Counterexamples Assessment.

Score	Pre-assessment (No. of students scoring at level)	Post-assessment (No. of students scoring at level)
Level 3	0	1
Level 2	1	2
Level 1	1	4
Level 0	13	8

Student L's response aligned with arguments for an alternative, logically equivalent form of the original claim, namely, if $a \neq 0$ and $ab = 0$, then $b = 0$, an alternative form of conditional claims with "or" in the conclusion. Yet, this interpretation seemed unlikely. Neither [Project 1] nor the ECE intervention explicitly taught this alternative form, and this form is not a version of the contrapositive. Student L made no meta-level comments about an alternative form for the claim, and it seemed unlikely Student L would invent this form on her/his own. Instead, I interpreted Student L's response as involving pragmatic reasoning. This was because her/his response relied on the phrase "only if" as she/he discussed what must occur if one of the options in the conclusion, $a = 0$ or $b = 0$, was not met. I also interpreted Student L's response as involving model-based reasoning because she/he represented a variety of scenarios, including alternatives to the claim's conclusion. Student L's prose also suggested that she/he reasoned from these representations. Although Student L's reasoning was not contrapositive reasoning, the reasoning involved a consideration of what must occur if an alternative to part of the conclusion were to occur. This consideration seemed leverageable toward contraposition reasoning as discussed by Antonini (2004), consideration of what must occur if the conclusion of a general claim was not met.

On the post-assessment, the decrease in converse errors, from nine to five, explained much of the improvements in students' responses. Student M's response on the post-assessment was representative of students who made these improvements:

Develop a viable argument: "Counter: ab or $ab = 0$ but a and b are greater than or less than 0. $0x = 0$. $1x$=infinity or greater. Anything multiplied by zero will equal zero; any number but zero multiplied by any number but zero will not equal zero. This claim is true." Argument viable? "I believe it is viable because I show what a counterexample is and I think I showed that it is not possible." (Student M, post-assessment response to Item 1, Level 3, counterexample description, pragmatic reasoning, and exploration of alternative models)

Student M's statement, "Any number but zero multiplied by any number but zero will not equal zero," illustrated the reasoning that I coded as contrapositive-type reasoning. In context, "any number but zero" was taken to mean "any number not equal to zero."

Student M followed an approach similar to that taught in the ECE intervention. Student M described the possible counterexamples to the original claim and argued that examples with these properties do not exist. However, Student M's post-assessment response lacked explicit discussion of a method/mode of argumentation.

Yet, Student M's statements about constructing descriptions of counterexamples and showing they do not exist suggested that Student M understood that her/his approach applied to any conditional claim. These statements were taken as sufficient to warrant a score of Level 3.

On the post assessment, four other students constructed responses for Item 1 that aligned with the contrapositive structure $\neg q \rightarrow \neg p$,arguing that when $a \neq 0$ and $b \neq 0$, $ab \neq 0$. However, all those who constructed contrapositive-type arguments also included examples of conforming cases. The following response from Student I is representative:

Develop a viable argument: "Claim: If $ab = 0$, then $a = 0$ or $b = 0$. Foundation: $a = 4,\ b = 3,\ 4{\bullet}3 = 12$[;] $a = 9,\ b = 2,\ 9{\bullet}2 = 18$ [;] $a = 4;\ b = 0,\ 4{\bullet}0 = 0$ [Student included three more conforming cases.]. Narrative link: If you multiply any counting number by zero, the product of that equation will be equal to zero, but if you multiply two counting numbers by each other, the product will equal more that [sic] zero." Argument viable? "Yes, my argument is viable because I have eliminated the possibility of counterexamples and addressed the conditions and conclusion." (Student I, Level 2, post-assessment response, contrapositive-type structure)

The inclusion of this unnecessary or irrelevant information about conforming cases suggested that the students were not following a strict procedure for contrapositive argumentation/proving. Instead, these students appeared to be reasoning pragmatically, discussing what would happen if the conclusion of the claim were not satisfied. Yet, the inclusion of conforming cases made it plausible that converse errors were still present on the post-assessment. In fact, all students who gave contrapositive-type arguments noted that 0 times any number is 0.

Finally, students who constructed contrapositive-type arguments, with the exception of Student M, provided no or insufficient meta-level justifications for the validity of the approach. As seen in Student I's response, above, the student reported eliminating counterexamples but offered no details about what counterexamples might look like or how his/her argument eliminated them.

On the post-assessment, three of the 15 students responded to Item 1 with arguments that addressed what would happen if one of the variables in the conclusion were not 0, similar to the pre-assessment response from Student L presented earlier. Student L's post-assessment response still contained this form of reasoning and was representative of the others:

Develop a viable argument: "Claim: If ab= 0, then a= 0 or b= 0. Foundation: Assume ab= 0. For ab to equal 0, either a or b must be equal 0. [Student draws two rectangles, each with "0" in the interior. One rectangle has one leg labeled b ; the other rectangle is labeled with "a not equal 0."] Narrative link: For ab to equal 0, either a or b must equal 0. This is because the only way to multiply a number that does not equal 0 with another number and have them equal 0 is if the other number equals 0. This means the only way ab= 0 is if a= 0 or b= 0." Argument viable? "In [sic] think my argument is viable because it uses an indirect argument to prove that there can't be counterexamples to the claim. If there can't be counterexamples, then the claim must be true." (Student L, post assessment response, Level 2, pragmatic reasoning and perhaps reasoning with models).

As seen above, the reasoning in Student L's arguments changed little from pre to post.

The only substantive difference in Student L's pre/post assessment responses to Item 1 was she/he made meta-level comments about her/his argument on the post-assessment, which were lacking in her/his pre-assessment response. On the post-assessment, student L asserted that 1) her/his argument was indirect (which it is not), 2) that the argument eliminated counterexamples, and 3) that general claims are true if there can't be counterexamples. Yet, these comments were not specific enough to the context of the argument to warrant a score higher than Level 2.

Finally, on both the pre- and the post-assessments, several students misinterpreted the item or gave responses that appeared irrelevant, all scoring at Level 0. One of these students mentioned eliminating counterexamples but did not offer any specific

Table 5
Item 3 Categorical Data, Pre/Post-Eliminating Counterexamples Assessment.

Score	Pre-assessment (No. of students scoring at level)	Post-assessment (No. of students scoring at level)
Level 3	0	1
Level 2	0	3
Level 1	0	3
Level 0	15	8

reference to the context of the claim presented. Because context-level details were absent, it was unclear whether the phrase "eliminating counterexamples" was meaningful to the student, or whether "eliminating counterexamples" was merely a phrase the student learned to use during the intervention.

6.2. Changes in students' critiques of indirect arguments

Items 3 and 5 presented students with indirect arguments for generalizations. On the pre-assessment, students tended not to recognize the indirect arguments as viable. Most students critiqued the indirect arguments for not addressing the objects described in the claims' conditions.

On the post-assessment, students' scores improved on these indirect argument items (see Tables 5 & 6). More students (7 verses 0 on the pre-assessment) recognized the method/mode of argumentation as indirect, and many students (4 verses 0 on the pre-assessment) noted that the argument demonstrated that counterexamples to the generalization were impossible.

On Item 3, the greatest differences among students' responses pre/post-ECE were in whether students recognized that the argument presented was an indirect argument. On the pre-assessment, all 15 students scored at Level 0. Five of these students critiqued the argument as not viable because it did not address the objects in the original claim. Students said things like "the claim is talking about squaring even numbers but in the support it only talks about squaring odd counting numbers without explaining how it connects to the claim" (Student L, pre-assessment response to Item 3).

On the pre-assessment, in response to Item 3, three students expressed structural representation or structural relation barriers, meaning the representations or structural implications in the argument were not understood. One student asked what the k stood for. Three other students presented responses irrelevant to the claim or responses so unclear the responses were impossible to classify. Three students left the item blank or indicated they had no idea how to respond. Another student critiqued the argument for not labeling a claim, foundation, and narrative link, a superficial critique scoring at Level 0.

On the post-assessment, the four students who scored at Level 2 or 3 on Item 3 either noted that the argument eliminated counterexamples or that the argument argued for the contrapositive, or both, and elaborated on the argument approach in the context of the task. Student L and Student N's responses were representative:

Viable? "This is a viable argument because it proves that if you don't meet the conclusion you couldn't have the conditions. This proves that there can't be a counterexample." Method? "The arguer is using an indirect argument. I think they are using an indirect argument because they are arguing for the contrapositive of the claim (If n^2 is odd, then n is odd) instead of directly arguing for the original claim." Improve? "It would make the argument more viable if they explained how this connects to/proves the claim." (Student L, post-assessment response to Item 3, Level 3, "Contrapositive" and "indirect" used correctly in context)

Viable? "Yes, because the arguer explains how squaring an odd number would result in an even number, plus one, making the result odd." Method? "An indirect argument, because they argue for the opposite of the claim to therefore prove the claim." Improve? "Use the same method of argumentation in the actual claim as they did in the support." (Student N, post-assessment response to Item 3, Level 2, indirect argument; correct content-level discussion).

Student L and Student N's responses were in stark contrast to their pre-assessment responses to Item 3. As mentioned above, on the pre-assessment, Student L asserted that the claim and support were mismatched, and Student N wrote, "No [the argument is not viable], because they used a counterexample when they used n as an odd number." Student N's post-assessment response to Item 3 also demonstrated improvement because the student recognized the indirect approach and discussed it in context.

Student N's post-assessment response lacked meta-level justifications for the indirect approach and the students response to the prompt "Improved?" suggested a preference for a direct argument, an argument where the support and claim aligned. Together these features raised questions about how well the student understood the validity of indirect argumentation/proving.

Table 6
Item 5 Categorical Data, Pre/Post-Eliminating Counterexamples Assessment.

Score	Pre-assessment (No. of students scoring at level)	Post-assessment (No. of students scoring at level)
Level 3	0	1
Level 2	0	5
Level 1	0	2
Level 0	15	7

On the other hand, Student L's post-assessment response to Item 3 was particularly powerful because the student articulated a meta-level understanding of indirect argument, explaining how counterexamples were eliminated. Student L also discussed the approach in the context of the objects in the claim.

Most of the students who scored at Levels 0 or 1 on Item 3 on post-assessment did mention counterexamples and/or eliminating counterexamples (9 of 11) but did not express a rich understanding of these notions and how these notions related to the argument presented. Two students gave responses on the post-assessment that suggested that the ECE intervention was helpful to them in recognizing indirect arguments, but at the same time, their responses suggested that their understanding of the validity of an indirect argument approach was tentative.

For example, on the post-assessment, Student A recognized the contrapositive structure in the argument presented in Item 3, and she/he acknowledged that the original claim and its contrapositive have the same counterexamples. However, Student A expressed uncertainty about the argument's viability. In fact, on the post-assessment, Student A maintained a view that the argument presented in Item 3 was a mismatch for the claim.

Viable? "[Crossed out as in original] because their support argues for a different claim: which is if n isn't even, then n^2 isn't even. … this could also be the contrapositive because its [*sic*] switched and negated and they have the same counterexamples." Method? "I think they are using indirect argument because their support supports the claim's contrapositive." Improve? "Explain why this supports it when the support shows odd numbers and the claim is about even numbers." (Student A. post-assessment response to Item 3, Level 2, acknowledged contrapositive reasoning and the equivalence of counterexamples for a claim and its contrapositive, but expressed uncertainty about contrapositive argumentation)

Student A correctly identified the conditional claim that the support argued for directly. However, because Student A expressed uncertainty about the validity of the method/mode of argumentation, the response did not score at Level 3.

Students made the largest categorical improvements pre/post in their responses to Item 5 of the assessment (see Table 6). The greatest differences pre/post-assessment were in whether or not students recognized the argument presented as an indirect argument.

On the pre-assessment, no student noted that the argument in Item 5 was an indirect argument. Five students said the claim and support did not match/agree. Five students labeled the argument as viable but provided little information about what features made it viable. Three of these students addressed the support's algebraic structures but made no statements about how the support connected to the claim.

On the post-assessment, no student critiqued the support in Item 5 for not addressing the mathematical objects in the claim. One student labeled the argument as "indirect" but scored at Level 0 because there was no meaningful discussion of why this label applied. In total, nine of 15 students' post-assessment responses correctly described the argument presented in Item 5 as an indirect argument, a contrapositive argument, or both (six/six/three students, respectively). Below are two students' pre/post responses that were representative of changes made, or not made, pre/post.

Viable? "This argument is not viable, because the support does not explain how it connects to the claim, and even though it says that if x = y, it does not prove why if $x^2 \neq y^2$then $x \neq y$." Method? "I don't know what method of argumentation this is." Viable? "They would need to add how it connects to the claim and why it proves the claim." (Student L, pre-assessment response to Item 5, Level 0, asserts that the support does not match the claim)

Viable? "I don't think this argument is a viable argument. I don't think it's a viable argument because it does not prove that the claim will be true for all cases." Method? "I think the argument is using an indirect argument. I think they are using an indirect argument instead of arguing for the original claim, they are arguing for the contrapositive of the claim." Improve? "Explain why x^2 must equal y^2 if $x = y$." (Student L, post-assessment response to Item 5, Level 1, indirect mode recognized but other information casts some doubt about understanding)

As seen above, on the pre-assessment, Student L critiqued the support found in Item 5 for not addressing the mathematical objects in the claim. On the post-assessment, Student L distinguished the indirect approach in Item 5 from a direct approach, saying "indirect instead of arguing for the original claim." On the post-assessment, Student L also critiqued the argument in Item 5 for not having sufficient context-level details. This was a fair critique of the argument given its purely symbolic representation. Yet, Student L's post-assessment response did not justify the indirect mode of argumentation and lacked specific context-level details, so her score could not rise above Level 1.

Student N gave a similar response to Item 5 on the pre-assessment but constructed a stronger response of the post-assessment:

Viable? "No, because the support statement does not actually agree with the claim." Method? "I don't have a clue to be honest. I'm sorry." Improve? "Use a support that agrees with the claim." (Student N, pre-assessment response to Item 5, Level 0, asserts support does not match the claim.)

Viable? "Yes, because in both instances, the claim and the support, x can work for both positive and negative numbers, therefore making the claim viable for all real numbers." Method? "An indirect argument, because they introduce the contrapositive, therefore arguing for contrapositive indirectly argues for the claim." Viable? "Show examples of how it applies to all real numbers." (Student N, post-assessment response to Item 5, Level 2, contrapositive argument associated with indirect argument; generality noted)

Similar to Student L's post-assessment response to Item 5, Student N distinguished the argument presented in Item 5 from a direct argument and correctly identified the argument as a contrapositive argument. Yet, Student N failed to justify the approach as a method for eliminating counterexamples.

On the post-assessment, two students used terms and phrases that were presented in the intervention but did not include sufficient information to warrant a score above Level 0 or 1. For example, Student M, who scored at Level 0 on the post-assessment, called for "counterexample elimination" in response to the prompt about what could be improved in the argument. Yet, Student M failed to address whether the argument presented in Item 5 attempted to eliminate counterexamples. Student I, who scored at Level 1 and

Table 7
Item 2 Categorical Data, Pre/Post-Eliminating Counterexamples Assessment.

Score	Pre-assessment (No. of students scoring at level)	Post-assessment (No. of students scoring at level)
Level 3	2	2
Level 2	1	5
Level 1	2	4
Level 0	10	4

labeled the argument as indirect, mentioned that *not the conclusion* was addressed in the argument but did not mention that *not the conditions* was also addressed. This omission casted serious doubts about the student's understanding of indirect argument. Perhaps this student simply associated the term "indirect argument" with arguments addressing *not the conclusion* of a conditional claim.

Three post-assessment responses to Item 5 were too vague to categorize. One student gave an instantiation conforming to the support given in Item 5 but then wrote, "I don't know." Another student incorrectly described the argument in Item 5 as direct because the support "met the conditions and the conclusion." Both of these responses scored at Level 0.

6.3. *Critiquing direct arguments*

Items 2 and 4 presented students with direct arguments for generalizations. These items offered opportunities to assess students' ability to distinguish between direct and indirect arguments. Also, even though the ECE intervention focused on constructing and critiquing indirect arguments, the ECE framework worked well for direct arguments. As Tables 7 and 8 illustrate, students scored poorly on these items on the pre-assessment but made improvements on the post-assessment.

On the pre-assessment, no student described the argument presented in Item 2 as a direct argument. Responses from the 10 students who scored at Level 0 tended to be vague. The following was representative:

Viable? "No because the math is wrong." Method? "If then because he claimed if then." Improve? "Correct calculations." (Student M, pre-assessment response to Item 2, Level 0, vague; possible structural representation barriers)

Student M's response that the calculations were incorrect suggested possible structural representation and structural relation barriers to reading the argument. Two other students who scored at Level 0 indicated that they did not understand how $2n + 1$ and $2m + 1$ were used to represent odd numbers or how the calculations demonstrated that the conditions implied the conclusions. Thus, these students were not able to read the argument correctly because they struggled with the variable expressions and how they were used.

Among the students who appeared to understand part or all of the variable expressions and the implications on these expressions, one student, scoring at Level 1, noted that the variable used in the argument was a feature of viable argumentation. This student also noted that the distributive property was used.

The strongest response to Item 2 on the pre-assessment came from Student L, who scored a Level 3, and wrote that the argument used logical chains. She/he also discussed how the conditions were used to show the conclusion:

Viable? "This argument is viable. I think it's viable because in the support it applies to the conditions whilst also proving that the conclusion is true." Method? "The arguer is using logical chains as their method of argumentation. I think this because the arguer is saying that since a & $b$can be written as $2k + 1$ or $2n + 1$ and $2k + 1 + 2n + 1$ can be simplified down to $2(k + n + 1)$ then $a + b$ is even because $2(k + n + 1)$ will be even." Improve: "The arguer might want to give an example." (Student L, pre-assessment response to Item 2, Level 3, explanation of direct argument in context and in general)

Like Student L, another student who scored at Level 1, noted that the argument applied to all cases.

On the post-assessment, eight students explicitly described the argument in Item 2 as a direct argument. Some elaborated on this classification, discussing how the conditions implied the conclusion, and distinguished the argument approach in Item 2 from arguments for the contrapositive of a generalization. The following were representative of these types of responses:

Viable? "The argument is viable. I think this argument is viable because it proves that if you meet the conditions, you must meet the conclusion." Method? "The arguer is using a direct argument. I think they are using a direct argument because they are directly arguing for their claim, not arguing for the contrapositive of their claim." Improve? "Something that would make this argument more viable is if it had a narrative link that explained why $2(k + n + 1)$ must have an even outcome." (Student L, post-assessment response for Item 2, Level 2, direct; conditions imply conclusion)

Table 8
Item 4 Categorical Data, Pre/Post-Eliminating Counterexamples Assessment.

Score	Pre-assessment (No. of students scoring at level)	Post-assessment (No. of students scoring at level)
Level 3	0	0
Level 2	0	4
Level 1	1	3
Level 0	14	8

Viable? “Yes, because they state how an odd number is essentially an even number plus one, so if you add a and b together, you’ll always get a number divisible by two, making it even.” Method? “A direct argument, because they don’t introduce the contrapositive, and instead break down the problem as it is.” Improve? “If the arguer included an example with numbers, instead of just variables.” (Student N, post-assessment response to Item 2, Level 2, direct; key structure identified)

It was interesting that Student L’s score decreased from Level 3 to Level 2 pre/post despite the fact that the post-assessment response contained more features that aligned with what was taught in the ECE intervention than the pre-assessment response, such as contrasting direct and indirect argument approaches. This drop may have been due to the fact that a score of Level 3 required 1) context-based evaluations of the argument, 2) identification of argument method, and 3) meta-level justifications for the method identified. Perhaps, in response to the intervention, Student L chose to focus more on identifying the argument method and justifying the method than on evaluating context-level details. Consequently, Student L’s decreased score may not reflect negatively on the impact of the intervention.

In contrast to Student L’s response, other students who score above Level 0 on the post-assessment included elaborations on the context-based structures in the argument. For example, Student N, whose post-assessment response is displayed above, referred to the structure of “odd” in her/his critique of the argument.

Student N’s Item 2 post-assessment was considerably stronger that her/his pre-assessment response. On the pre-assessment, Student N scored at Level 0 and merely noted that the claim was general: “[The method is a] for all claim, because their claim applies to all examples of two odd counting numbers added together.” Students who scored above Level 0 on the post-assessment tended to be more explicit about the context-based structures in the argument, describing how the extra ones in any pair of odd numbers add to an even number.

In their responses to Item 4, students made improvements pre/post-assessment, although these improvements were minor for most students. The limited improvements may have been due to structural representation and structural relational barriers, particularly when addressing the implication “$x < 0$ implies $-x > 0$.” Fourteen of 15 students made comments on pre-assessment that suggested these barriers. Below is a representative example:

Viable? “I’m not sure because they say $-x$ is positive.” Method? “A counterexample because if they are proving the claim by saying something else.” Improve? “Maybe add the word assume or give a visual to help understand.” (Student A, pre-assessment response to Item 2, Level 0, unsure about the role of $-x$.)

Student A’s comments, “proving … by saying something else,” suggest that the student viewed the argument as indirect, and yet there is no evidence that the student viewed the argument approach as a viable method/mode in general. On the pre-assessment, two other students misinterpreted the support as not directly addressing the claim, commenting that the argument is not viable because the support and claim do not match.

In fact, on the pre-assessment, no student described the argument in Item 4 as a direct argument. One student incorrectly labeled the argument as “indirect.” Four students offered no response to the “method?” prompt. Another four commented on superficial features (e.g., critiquing the argument for not labeling a claim, foundation, or narrative link) or presented comments too vague to categorize. Only one student’s response on the pre-assessment contained comments suggesting that the argument’s direction was followed. That student said, “Yes [viable], because $-x$ could be a negative, less than 0, a negative times negative = positive > 0,” and yet, this student’s response contained the troubling statement, “$-x$ could be a negative,” which calls into question whether the implication $x < 0$ implies $- x > 0$ was understood.

On the post-assessment, seven students described the argument presented in Item 4 as a direct argument. Five of these students elaborated on the argument in a manner suggesting that they understood what made the argument direct. The strongest response was from Student F:

Viable? “Because the only way that x could be less than zero is if it is a negative number and a negative number multiplied by itself is a positive number, because a negative multiplied by a negative is positive.” Method? “Direct because they say if the conditions (if $x < 0$) then the conclusion ($x^2 > 0$).” Improve? “Nothing this argument is very viable.” (Student F, post-assessment response to Item 4, Level 2, direct argument approach explained)

Student F elaborated on the key implication “$x < 0$ implies $-x > 0$” in a manner suggesting that the implication was understood. The converse reasoning suggested by the phrase “the only way” was problematic but did not appear significant relative to the student’s overarching response. After all, Student F elaborated on how the direct-argument label was assigned. Yet, this elaboration did not express a justification for the method/mode’s validity, so a Level 3 score was not assigned.

On the post-assessment, five students mentioned some facet of the ECE approach to viable argumentation. Three of these five called for some facet of the approach to be explicitly included in the argument, but these three students did not acknowledge that the argument already attempted to eliminate counterexamples. The following is representative:

Viable? “Yes because it shows why the claim is true.” Method? “Generalization because it states any number will work.” Improve? “Showing whether or not a counterexample exists.” (Student M, post-assessment response to Item 4, Level 0, calls for addressing whether or not counterexamples exist)

Some improvement were due to students contrasting the argument method used in Item 4 with a contrapositive argument. The following is an example:

Viable? “No, because they state that $-x > 0$, contradicting what they said, that if $x < 0$. Also, they claimed $-x$ is a positive number, which is absurd, considering you can’t have a positive, negative number.” Method? “A direct argument, because they don’t introduce a contrapositive, therefore they must be arguing for the claim directly.” Improve? “Include viable information within their claim.” (Student N, post-assessment response to Item 4, Level 1, direct argumentation contrasted with contrapositive argumentation)

Student N did not elaborate on the mode of argumentation used in Item 4, nor did she/he elaborate on the algebraic transformations found in the argument. Yet, Student N contrasted the argument approach found in Item 4 with arguments for the contrapositive, which at first glance made her/his post-assessment response appear superior to her/his pre-assessment response, where she/he wrote, “I truly have no idea what to do.” However, on the post-assessment, Student N also wrote two statements—“they claim that $-x$ is a positive number, which is absurd,” and “contradicting what they said [earlier]”—which created doubts about whether Student N followed the mode of argumentation beyond noting that the argument started with the conditions and ended with the conclusion. These doubt-raising comments prevented her/his post-assessment score from rising above a Level 1.

Structural representation/relation barriers remained on the post-assessment. Some students improved their scores by identifying the argument approach correctly as a direct argument but still expressed these barriers. Student L’s response below was an example of these issues. Student L appeared to follow much of the argument, but her/his request for more explanation for why $-x$ is positive suggested that this implication was not understood:

Viable? “I think the argument is viable. I think it is viable because it proves that no matter what x equals, x^2 will always be greater than 0.” Method? “The arguer is using a direct argument. I think they are using a direct argument because the foundation is directly arguing for the original claim.” Improve? “It would make it more viable if they explained why $-x$ must be a positive number.” (Student L, post-assessment response to Item 4, Level 2, direct argument label; structure representation and relation barriers)

6.4. Summary of results

As noted in previous sections, on the pre-assessment, over a third of students critiqued the contrapositive arguments for not addressing the original claim or not addressing the mathematical objects in the claim, and no student recognized the indirect arguments presented as using a valid method/mode of proof. The pre-assessment results were surprising, given that the students had received instruction on contrapositive proving/argumentation (and the ECE approach, briefly) prior to the assessment.

After the intervention, over half the students correctly distinguished direct arguments from indirect arguments, labeling the indirect arguments correctly as indirect, contrapositive, or both. On the post-assessment, more than a third of the students made comments contrasting indirect arguments and direct arguments in response to at least one of the three prompts on at least one item. These students elaborated on the approach sufficiently to conclude that the students understood how indirect arguments/proofs eliminate the possibility of counterexamples to a generalization.

The current study also provided evidence that students’ abilities construct an contrapositive-like argument can be improved with sufficient experience describing counterexamples to generalizations. In response to pre-assessment Item 1, only two students addressed alternatives to at least part of the conclusion of the claim. No student constructed a contrapositive argument, and nearly all students made a converse error.

On the post-assessment, over half the students constructed an argument that had features of indirect argumentation/reasoning, such as considering what occurs when one part or all the conclusion is not met. Students either argued that “not the conclusion implies the conditions do not occur” or constructed an argument for a logically equivalent form, $ab = 0$ and $a \neq 0$ implies $b = 0$. Argument approaches for alternative forms of conditional claims, such as those with “or” in the conclusion, were not introduced in [Project 1]/ECE interventions. It is plausible that students whose responses appeared to address an alternative form were in fact considering cases of not the conclusion to the original claim.

Overall, only a few students explicitly constructed descriptions of possible counterexamples to the claim in Item 1, but about half of the students discussed whether their argument eliminated counterexamples. These findings support an assertion that the ECE approach can improve at least some Grade 8 students’ ability to construct contrapositive arguments.

Finally, the current study also provides evidence that students’ abilities construct meta-level validations of contrapositive arguments can be improved with sufficient experience describing counterexamples to generalizations. On the pre-assessment, viable validations of indirect arguments were nearly absent, and only few students made meta-level comments about the direct arguments presented. On post-assessment, approximately one-third of the students gave responses with meta-level comments that justified contrapositive approaches to at least one item. One negative finding came from the student who critiqued every argument for not explicitly using the ECE approach, an unintended effect of the intervention.

7. Discussion

Students’ difficulties with contrapositive proof/proving are well documented. (e.g., Antonini & Mariotti, 2008; Yopp, 2017). Yet research is lacking on classroom-based interventions for improving students’ ability to recognize, use, understand, and validate contrapositive arguments/proofs and to contrast the methods/modes with other methods of proof (Stylianides et al., 2017). Studies of interventions that improve students’ understanding and use of proof/proving methods are needed (Stylianides et al., 2017).

The current study provides evidence that students’ abilities to judge the direction of an argument can be improved with sufficient experience describing counterexamples to generalizations. Yet, it should be noted that most of the students whose responses improved pre/post assessment did not use the ECE approach explicitly or verbatim. It is not known whether these students mentally considered descriptions of possible counterexamples and reflected on whether examples satisfying these descriptions were eliminated. It is possible that students used the approach mentally without recording their mental activities on the assessment. However, it is also possible that the ECE approach helped students learn about contrapositive argumentation and its validity but did not become a strategy for assessing or constructing arguments. Either way, experiences with the ECE approach resulted in improved scores on indirect, and direct, argument items.

Distinguishing a direct argument from an indirect argument can be rather rote. A student might correctly identify an argument as an indirect because the argument addresses objects that are "not the conclusion" of the claim. Yet it should not be overlooked that students in this study who scored above Level 1 often went beyond simply distinguishing contrapositive arguments from direct arguments. These students applied criteria for judging whether a contrapositive argument was viable, such as whether all cases were addressed and whether counterexamples were eliminated, and at times, students explained why the argument eliminated counterexamples in the context of the claim.

This discussion brings to the forefront the question of how much experience with indirect reasoning activities is needed for eighth-grade students to improve their contrapositive reasoning. The ECE intervention included more activities involving indirect argument and the ECE justification scheme than the [Project 1] intervention. During the ECE intervention, students developed numerous descriptions of possible counterexamples to conditional claims and descriptions of possible counterexamples to the original claim and its contrapositive. Students compared the descriptions and the sets the description determined. During [Project 1] these activities only occurred in a few places, where the content standards called for a topic that aligned well with indirect argument (e.g., proving the converse of the Pythagorean Theorem).

One finding that may be important to future researchers is the association between student improvements in response to the intervention and student past mathematics achievement. Students who saw the best improvements were already proficient or advanced in mathematics, which may explain some of the success. In general, student with low-levels of mathematics achievement saw little or no improvement in response to the intervention.

Future research might address why low-achieving students tended not to improve knowledge, practice, and skill in response to an intervention. As noted in the findings, I observed that the low-achieving students expressed difficulties with the content of the tasks, including difficulties understanding the mathematical objects in the claims and their support and difficulties navigating the representations and transformations on these representations. Moreover, informally, and perhaps anecdotally, I noted that the low-achieving students had difficulty maintain focus and engagement during the intervention and were resistant to writing explicit descriptions of counterexamples using the connectives "and" and "but," which were encouraged during the intervention when describing the possible counterexamples to a claim. It is possible that proficient or advanced status, as determined by mathematics achievement assessments, is a proxy for levels of engagement.

8. Limitations

This was a case study of one class in one special learning environment, not generalizable to other learning environments. Like other case studies, it demonstrates possibilities for future research in different contexts.

Another limitation is that the same assessment was used pre- and post-ECE. No validated assessments existed, and the sample was insufficiently large to equate pre- and post-ECE scores on two different assessments. The use of the same assessment might explain some students' improvements.

9. Conclusion

In response to the research question—In what ways do a group of students enrolled in eighth-grade mathematics (U.S. system, age 13) change their contrapositive reasoning in response to the ECE intervention?—I found that, post-intervention, students improved their ability to distinguish indirect arguments from direct arguments, and many students acknowledged the validity of indirect arguments. Some of these students could communicate meta-level justifications for the approach. Barring a few exceptions, proficient or advanced students made progress toward the intervention's goals. Post-assessment, more students constructed a viable indirect argument for a generalization, than on the pre-assessment, and more students correctly identified direct and indirect arguments. However, only one student explicitly used the ECE approach to construct his/her argument. Others may have developed descriptions of possible counterexamples mentally when developing their arguments but did not explicate the approach.

I found that over half the students used the ECE standard for proof to critique at least one argument presented on the post-assessment. Several used the standard effectively to critique both direct and indirect arguments. Yet students who presented responses on the post-assessment at levels above 0 did not always use the ECE approach explicitly. While, an explicit description of the collection of all possible counterexamples was key to ECE approach, these explicit descriptions were not always found on high-level responses. Perhaps, the ECE intervention was effective in getting many students to use and validate contrapositive approaches, but the explicit use of the ECE approach was not viewed as critical by students as they constructed or critiqued arguments.

Students' past and concurrent mathematics achievement was associated with their use and understanding of the ECE approach and indirect arguments on the assessments. Low-achieving students did not use the approach, or did not use it correctly, and did not improve their scores, but they also appeared to have difficulty negotiating intervention tasks and assessment items in general. It is unknown whether the ECE approach was accessible and useful to these students, because too often low achieving students' activities were confounded by errors and misunderstandings unrelated to the ECE approach. The other low-achieving students in the larger group made comments during the intervention and on the assessments expressing difficulty understanding the content of the claims and the algebraic manipulations in the claims' supports. Future studies might explore ways to improve students' use of the ECE approach as an explicit strategy for developing or critiquing an argument in design studies. Other studies might explore the ECE approach with lower-achieving students in mathematical contexts more accessible to them.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. DRL 1317034. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jmathb.2020.100794.

References

Alibert, D., & Thomas, M. (2002). Research on mathematical proof. In D. Tall (Vol. Ed.), *Advanced mathematical thinking: Vol. 11*, (pp. 215–230). Dordrecht, Netherlands: Springer.

Antonini, S. (2003). Non-examples and proof by contradiction. In N. A. Pateman, B. J. Dougherty, & J. Zilliox (Vol. Eds.), *Proceedings of the 27th Joint Meeting of PME and PMENA: Vol. 2*, (pp. 49–55).

Antonini, S. (2004). A statement, the contrapositive and the inverse: Intuition and argumentation. In M. Johnsen Høines, & A. Berit Fuglestad (Vol. Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education: Vol. 2*, (pp. 47–54).

Antonini, S., & Mariotti, M. A. (2008). Indirect proof: What is specific to this way of proving? *ZDM–The International Journal on Mathematics Education, 40*, 401–412. https://doi.org/10.1007/s11858-008-0091-2.

Brown, S. A. (2014). On skepticism and its role in the development of proof in the classroom. *Educational Studies in Mathematics, 86*, 311–335. https://doi.org/10.1007/s10649-014-9544-4.

Brown, S. A. (2018). Are indirect proofs less convincing? A study of students' comparative assessments. *Journal of Mathematical Behavior, 49*, 1–23. https://doi.org/10.1016/j.jmathb.2016.12.010.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, 17*, 391–416. https://doi.org/10.1016/0010-0285(85)90014-3.

Dumas, B. A., & McCarthy, J. E. (2007). *Transition to higher mathematics: Structure and proof.* New York, NY: McGraw-Hill.

Esty, W. W., & Esty, N. C. (2013). *Proof: Introduction to higher mathematics*Retrieved from(7th ed.). http://estymath.com/Proof.html.

Fan, L., Qi, C., Liu, X., Wang, Y., & Lin, M. (2017). Does a transformation approach improve students' ability in constructing auxiliary lines for solving geometric problems? An intervention-based study with two Chinese classrooms. *Educational Studies in Mathematics, 96*, 229–248. https://doi.org/10.1007/s10649-017-9772-5.

Freudenthal, H. (1973). *Mathematics as an educational task.* Dordrecht, Netherlands: Reidel.

Goos, M. (2004). Learning mathematics in a classroom community of inquiry. *Journal for Research in Mathematics Education, 35*, 258–291. https://doi.org/10.2307/30034810.

Harel, G. (2001). The development of mathematical induction as a proof scheme: A model for DNR-based instruction. In S. Campbell, & R. Zazkis (Eds.). *Learning and teaching number theory: Research in cognition and instruction* (pp. 184–212). Westport, CT: Greenwood Press.

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

Jahnke, R. N., & Wambach, R. (2013). Understanding what a proof is: A classroom-based approach. *ZDM–The International Journal on Mathematics Education, 45*, 469–482. https://doi.org/10.1007/s11858-013-0502-x.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge, United Kingdom: Cambridge University Press.

Komatsu, K. (2017). Fostering empirical examination after proof construction in secondary school geometry. *Educational Studies in Mathematics, 96*, 129–144. https://doi.org/10.1007/s10649-016-9731-6.

Krantz, S. G. (2002). *The elements of advanced mathematics* (2nd ed.). Boca Raton, FL: CRC Press.

Lakatos, I. (1976). *Proofs and refutations: The logic of mathematical discovery.* Cambridge, United Kingdom: Cambridge University Press.

Leron, U. (1985). A direct approach to indirect proofs. *Educational Studies in Mathematics, 16*, 321–325. https://doi.org/10.1007/BF00776741.

Mariotti, M. A. (2013). Introducing students to geometric theorems: How the teacher can exploit the semiotic potential of a DGS. *ZDM–The International Journal on Mathematics Education, 45*, 441–452. https://doi.org/10.1007/s11858-013-0495-5.

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

O'Brien, T. C. (1972). Logical thinking in adolescents. *Educational Studies in Mathematics, 4*, 401–428. https://doi.org/10.1007/BF00566998.

Reid, D., & Dobbin, J. (1998). Why is proof by contradiction difficult? In A. Olivier, & K. Newstead (Vol. Eds.), *Proceedings of the 22th Conference of the International Group for the Psychology of Mathematics Education: Vol. 4*, (pp. 41–48).

Sandefur, J., Mason, J., Stylianides, G. J., & Watson, A. (2013). Generating and using examples in the proving process. *Educational Studies in Mathematics, 83*, 323–340. https://doi.org/10.1007/s10649-012-9459-x.

Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education, 26*, 114–145. https://doi.org/10.2307/749205.

Stylianides, G. J., & Stylianides, A. J. (2009). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education, 40*, 314–352. Retrieved from https://www.nctm.org/publications/journal-for-research-in-mathematics-education/.

Stylianides, G. J., Stylianides, A. J., & Weber, K. (2017). Research on the teaching and learning of proof: Taking stock and moving forward. In J. Cai (Ed.). *Compendium for research in mathematics education* (pp. 237–266). Reston, VA: National Council of Teachers of Mathematics.

Yopp, D. A. (2011). How some research mathematicians and statisticians use proof in undergraduate mathematics. *The Journal of Mathematical Behavior, 30*(2), 115–130.

Yopp, D. A. (2014). Undergraduates' use of examples in online discussions. *The Journal of Mathematical Behavior, 33*, 180–191. https://doi.org/10.1016/j.jmathb.2013.11.004.

Yopp, D. A. (2015). Prospective elementary teachers' claiming in responses to false generalizations. *The Journal of Mathematical Behavior, 39*, 79–99. https://doi.org/10.1016/j.jmathb.2015.06.003.

Yopp, D. A. (2017). Eliminating counterexamples: A Grade 8 student's learning trajectory for contrapositive proving. *The Journal of Mathematical Behavior, 45*, 150–166. https://doi.org/10.1016/j.jmathb.2017.01.003.

Yopp, D. A., & Ely, R. (2016). When does an argument use a generic example? *Educational Studies in Mathematics, 91*(1), 37–53. https://doi.org/10.1007/s10649-015-9633-z.