The Mathematics Enthusiast

Manuscript 1636

On Convincing Power of Counterexamples

Orly Buchbinder

Rina Zazkis

On Convincing Power of Counterexamples

Orly Buchbinder[1], Rina Zazkis[2]
[1]University of New Hampshire [2]Simon Fraser University

Abstract: Despite plethora of research that attends to the convincing power of different types of proofs, research related to the convincing power of counterexamples is rather slim. In this paper we examine how prospective and practicing secondary school mathematics teachers respond to different types of counterexamples. The counterexamples were presented as products of students' arguments, and the participants were asked to evaluate their correctness and comment on them. The counterexamples varied according to mathematical topic: algebra or geometry, and their explicitness. However, as we analyzed the data, we discovered that these distinctions were insufficient to explain why teachers accepted some counterexamples, but rejected others, with seemingly similar features. As we analyze the participants' perceived transparency of different counterexamples, we employ various theoretical approaches that can advance our understanding of teachers' conceptions of conviction with respect to counterexamples.
Key words: Counterexamples, Proof, Conviction, Secondary Teachers

1 Introduction

Mathematically, one counterexample is sufficient as a proof that a claim is false. However, do different counterexamples have the same convincing power in disproving a conjecture? Our short answer is *no*, and we present a long elaboration to support this answer.

The question about convincing power of proof has been at the crux of the long-standing discussions among mathematicians and mathematics educators. Considering the roles of proof in mathematics, Hanna (2000) asserted that mathematicians prefer proofs that explain to proofs that "just prove", and that the main role of proof in mathematical classrooms is explanation, which is strongly related to the notion of conviction. Convincing and removing doubt in the truthfulness of a statement have been considered by many as important, even primary goals of proving (Brown, 2018; Harel & Sowder, 2007). Still others emphasize that achieving psychological conviction is not a necessary requirement of proof; there are convincing arguments that are not proofs, and valid deductive proofs that bestow little conviction. Some approaches suggest distinguishing between types of conviction: psychological feeling of certainty vs. logical certainty grounded in deductive reasoning (Harel & Weber, 2020;

[1] orly.buchbinder@unh.edu

Mamona-Downs & Downs, 2010). Weber and Mejia-Ramon (2015) pointed out to the ambiguity of the word conviction, suggesting the need to differentiate between different types of conviction, i.e., absolute and relative conviction. The absolute conviction refers to a person's feeling of complete removal of doubt, while relative conviction conveys that one's certainty in the probability that a statement is true exceeds certain threshold. We will return to this point later.

There is much less conceptualization for conviction in relation to counterexamples. Researchers have identified various challenges associated with creating, understanding, and using counterexamples, such as difficulty to recognize counterexamples, difficulty to accept counterexamples in disproving a general statement, a tendency to treat counterexamples as exceptions, or to seek multiple counterexamples (e.g., e.g., Buchbinder & Zaslavsky, 2019; Peled & Zaslavsky, 1997; Zazkis & Chernoff, 2008). Building on these ideas, we further examine the question of what makes a counterexample convincing. In this theoretical paper, we present two cases where prospective and practicing teachers assessed the legitimacy of proposed counterexamples and consider teachers' responses through a variety of theoretical lenses.

2 Two Illustrative Cases

Case 1: A quadrilateral with congruent diagonals.

Consider the following statement and a response from a student Dina:

Statement-A: Any quadrilateral with congruent diagonals has at least one pair of opposite congruent sides.

Dina's Response: I can draw two congruent segments (diagonals) that intersect each other, such that the resulting four parts are all of different lengths. Then, I connect the endpoints of the diagonals with segments, one after another, creating a quadrilateral. The resulting quadrilateral disproves the statement.

Statement-A is false. A counterexample to the Statement-A can be represented by a single figure (possibly with some measurements showing no additional special properties) or a description how to construct such a quadrilateral, or both. Dina provides a description of how to construct a general counterexample without including the figure. That is, following Dina's description, one can construct multiple counterexamples to Statement-A. When two diagonals intersect each other two pairs of vertical angles are created; the requirement that "the resulting four parts (of the diagonals) are all of different lengths" assures that the resulting pairs of triangles do not satisfy side-angle-side congruence condition, and therefore the resulting quadrilateral has no congruent opposite sides. In providing a description of how to construct a figure, Dina's counterexample is reminiscent of non-constructive proofs of existence, which "indirectly show a mathematical object exists without providing a specific example." (Courant & Robbins, 1996, p. 86).

In the ongoing research of the first author, the statement and the response were presented to different groups of practicing and prospective teachers. These teachers participated in a professional development or undergraduate course intended to refresh their knowledge of mathematical proof and connect it to proving in secondary geometry and algebra classrooms. All participants had an undergraduate degree in secondary mathematics education, or equivalent, and had completed at least two proof-based courses at the undergraduate level.

When presented with Statement-A and Dina's argument, only about a quarter of participants accepted Dina's solution as valid, by drawing a quadrilateral based on her description. One participant, who rated Dina's work highly explained:

"It shows the general counterexample of the above statement which is as clear as the single example given."

However, the overwhelming majority of the participants were critical of Dina's mathematical work. Some participants felt that without an accompanying sketch, the verbal explanation is insufficient.

"[If you say] You can, so just do it. You can make a picture instead of long describing. A drawing is more than words in this case."

"A drawing would provide support to the counterexample."

The fact that Dina did not construct a specific counterexample, but only described how to produce one elicited critique from participants, who wished to see a specific example supporting the argument. Several participants even questioned the validity of a counterexample that is not specific. One participant, who constructed a sketch based on Dina's description wrote:

"The instructions can be followed, and the quadrilateral produced has no pair of opposite congruent sides. Is it a counterexample if it is not specific?"

Others wrote similar comments, like: "This counterexample is not specific enough."
These responses and many similar ones, suggest that the participants had a preference towards a specific counterexample as opposed to a general one. This appears in discord with Peled & Zaslavsky's (1997) assertion that general counterexamples have greater explanatory power than specific ones, since they describe a class of counterexamples and, as such, shed light on why a claim is false. Similar to research on proof (e.g., Hanna, 2000) one can expect that being "explanatory" implies being convincing (although not the other way around). Yet most participants did not find Dina's description of how to construct a general counterexample convincing.
In light of these responses, we wondered whether there was a common tendency to prefer specific counterexamples over general counterexamples. We thus examined another task that juxtaposed specific and general counterexamples.

Case 2: Generating primes.

Consider the following statement and responses from two students:

Statement-B: If n is a natural number, then $n^2 + n + 17$ is prime.

Response from Yana's group: We substituted $n = 16$ into $n^2 + n + 17$ and got 289, which is no prime since $17 \times 17 = 289$. So the statement is false.

Response from Olga's group: We used algebra to see what happens if n is a multiple of 17. We substituted $n = 17k$ in $n^2 + n + 17$ and we got $(17k)^2 + 17k + 17 = 17(17k^2 + 1)$. This is a composite number, so the statement is false.

Statement-B is false. It can be disproved by a particular counterexample as suggested by Yana, or by a general counterexample which is a multiple of 17, as Olga suggested. However, there are also numerous counterexamples which are not multiples of 17, such as 16, 20 or 25.

Sometimes, general counterexamples may provide an insight for why a certain statement is false. This may allow for modifying the statements' domain and formulating a true statement on the reduced domain. However, modifying Statement-B by adding a condition that $n \neq 17k$ for $k \in \mathbb{N}$ does not result in a correct statement. Moreover, there seems to be no immediate way to modify the statement's domain to obtain a true statement.

In the ongoing research on the use of examples in proving, participants were asked first to determine if the statement is true or false, and then to determine the correctness of students' responses, to indicate their preference and to consider questions they would ask each student to support their mathematical thinking. We wondered whether the participants would show a clear preference towards a specific counterexample, similar to the case of Dina. Interestingly, the data did not support this conjecture. The participants' accepted both specific and general counterexamples, but their comments pointed to preference towards the latter.

The participants accepted Yana's specific counterexample as correct and legitimate disproof of the statement. At the same time, many comments expressed some sense of mistrust in Yana's understanding of the concept of a counterexample and sought supporting explanations. The participants also wanted to know how Yana came up with $n = 16$ being a counterexample, and whether additional counterexamples exist. Some representative comments are:

"I would tell the students that they have found a good counterexample and suggest that they explain why this is a counterexample. What about $n = 16$ makes the statement invalid? What other counterexamples can you find? Are there any relationships between the counterexamples?"

"I would ask if the student found any numbers that did hold this statement true. I would also ask if they were able to find a way to show that this expression is not always true without just getting lucky by plugging in 16."

In addition, participants indicated responding to Yana with a request to modify the statement "to make it true", as the following comment shows:

"I agree [with your argument] but are there some instances where this works? If so, can we make a general rule that satisfies those values of N?"

"Why does your counterexample prove that the statement in general is false? Is there a way you could change the statement to make it true?"

The approval of Yana's counterexample indicates that the participants were aware that a single counterexample is sufficient for disproving a general statement. This is contrary to the prevalent result in the literature showing that individuals tend to reject specific counterexamples as exceptions that do not disprove a general claim (e.g., Buchbinder & Zaslavsky, 2019). However, we also note that the participants seem unsatisfied with a single specific counterexample, wondering how the students came up with it, requesting more counterexamples, and expressing a wish for a more general counterexample.

Olga's argument appeared to satisfy these wishes. The participants' comments mainly focused on the algebraic work of Olga, as the following response shows:

"Using a counterexample is enough to prove something is false, but the student was able to actually use algebra to prove that the statement is false. They were able to generalize the fact that this statement is false."

Several participants were impressed with Olga's work, specifically pointing to the generality of the counterexample:

"The student found a valid, general counterexample and then supported it with correct algebra."

"Excellent counterexample, there is a general rule and not just a specific counterexample."

Reflecting on the two cases, we note that the generality of a counterexample was not accepted in Case 1 but praised in Case 2. Moreover, Yana's specific counterexample, while not disputed, was not fully embraced either. So, the overall preference towards either general or the particular counterexample cannot be clearly established. Hence, we wonder, what features of counterexample account for the participants' preferences? In what follows we address this question referring to a variety of theoretical constructs. We start with the apparent distinction between two cases: algebra vs. geometry, followed by the notion of example spaces that reflects the nature of school instruction. Then, we re-examine the general vs. particular connection, consider the notions of transparency and epistemic aims.

3 On Algebra vs. Geometry

The apparent difference between the two cases is that the statements are taken from different mathematical topics: geometry and algebra (number theory). Geometric objects are *figural concepts* (Fischbein, 1993) – they possess specific properties while intending to represent ideal conceptual properties. Chazan and Herbst (2011) emphasize that geometric figures, either sketched or rigorously constructed, are necessarily particular objects, characterized by certain dimensions, angle measures, side lengths and areas. Although geometric figures intend to convey generality and represent a whole class of geometric objects, each particular diagram contains superfluous visible features. Hence, geometrical objects presented as diagrams need to be unpacked, in order to be understood and used within the justification process. Learning to see the generality inherent in geometric figures is a complex process. For example, students have been observed to interpret diagrams accompanying proofs as specific objects and the theorems proven with the aid of such diagrams as applied only to those diagrams (Chazan, 1993). While the intended generality of geometric figures is assumed by mathematicians and mathematics teachers, students need to be enculturated into this view, in order to gain "the affordances of the particular to discuss the general, while minimizing the limitations that the particular has for discussing the general (Chazan and Herbst's, 2011; p. 10).

The nature of algebraic objects is quite different. The use of symbols in algebra allows one to represent an unknown number, or a whole class of numbers with a single letter, which in itself does not carry any characteristic features of that class of numbers. Further, the mathematical symbols themselves do not embody the particularity of the represented numbers in the way geometric figures do. In Chazan and Herbst's (2011) terms, "the affordance of literals (x's and y's), as opposed to numbers or geometrical diagrams, is that they aid in seeing generality without having to ignore some particularity" (p. 11). Therefore, the participants preference towards the algebraic counterexample of Olga may be attributed to the nature of algebraic symbols which carry greater generality. However, the generality of Dina's counterexample may be obscured not by the specific features of a particular figure, but by the absence of the figure itself. This may explain participants' request for a sketch, which would visualize both the property of congruent diagonals and the absence of additional special properties.

4 On Example Spaces

The notion of example space (Watson & Mason, 2005) – that is, a collection of examples that fulfil a particular function – may provide an additional explanation for the differences in participants' responses to counterexamples. Zazkis and Chernoff (2008) described several situations in which some counterexamples are more likely than others to be accepted by students. According to Mason (2006), "exemplariness resides not in the example, but in how the example is perceived" (p. 62). Echoing Mason's notion of "exemplary example" Zazkis and Chernoff discussed "what makes a counterexample exemplary". While any counterexample refutes a statement, only some are accepted by students with ease, while others are questioned or rejected as "exceptions". The authors introduced a notion of "pivotal (counter)example – an example that invites students to re-examine their prior claims, convictions, or initial intuitive inclinations.

Among their examples, Zazkis and Chernoff discussed the case of Selina, who initially considered 437, which is a product of 19 and 23, to be a prime number because it was a product of primes. When asked to examine her approach by considering numbers 6 or 15, both of which are also products of two primes, Selina did not change her opinion. However, the number 77, as a product of 7 and 11, served for Selina as a pivotal example that, "changed everything", that is, resulted in refuting the initial claim regarding 437 being prime.

Zazkis and Chernoff suggested that in order to serve a pedagogical purpose of rejecting a statement, a counterexample should fit within or be sufficiently close to the learner's *personal example space* – a construct describing an individual's repertoire of accessible examples and their construction (Watson & Mason, 2005). A counterexample that does not fit within a student's personal example space is likely to be questioned or rejected as an "exception."

In examining the expression $n^2 + n + 17$, the initial tendency for some participants was to check for primality of the result by substituting small numbers and observing that the expression generates prime numbers. The expression indeed generates prime numbers for the first 15 values of n. For some participants, this may result in an initial conviction that the statement is true. Both 16 and $17k$ are clearly within the participants' example spaces and either can be considered as a pivotal counterexample: 16 being a small enough number to test easily for generating a composite number, and $17k$ being simple enough algebraic expression to observe it generating a factorable expression. Thus, it is not surprising that both 16 and $17k$ were accepted by the participants as counterexamples and there was no apparent preference for either counterexample.

The notion of example space may also explain why Dina's construction protocol for a general quadrilateral with congruent diagonals was not embraced by the majority of participants: it may be outside the participants' personal example spaces. Special quadrilaterals (e.g., parallelograms, kites) represent a subset of all quadrilaterals, who may have no special properties or names. However, textbooks and curricula materials pay most attention to special quadrilaterals, while general quadrilaterals are seldom mentioned or discussed. Thus, in the context of school geometry, general quadrilaterals are outside the *instructional example space* which is "a set of collections of examples used by teachers and students as they work on a particular notion in an instructional context." (Biza, 2011, p. 130).

Beyond that, Dina's description does not fit the example space of how quadrilaterals are usually constructed. A common procedure for constructing quadrilaterals is to draw a quadrilateral first and then construct its diagonals. This is also a preferred mode of defining geometric objects, as individuals tend to favor definitions that rely on the salient features of geometrical objects, such as sides and angles, rather than features that require additional construction, such as diagonals, altitudes, or axes of symmetry (Zaslavsky & Shir, 2005). Dina's descriptive counterexample may fall outside the participants' instructional and personal example space for (a) being a general quadrilateral and (b) being defined procedurally based on the properties of diagonals.

5 On a General in the Particular and Vice Versa

In their seminal article, Mason and Pimm (1984) coined the notion of "seeing a general in the particular" with which they intended to "bring out some of the inherent difficulties in mathematical expressions of generality and their relation to the particular" (p. 287). Mason and Pimm discussed the example of the function $f(x) = |x|$ as a counterexample to the claim that continuous functions are differentiable. They noted that a lecturer presenting this example was seeing it as a generic example – a particular case which illuminates the general properties of the example to illustrate a general idea or a claim. However, students may have focused on a particular function without attending to its generality. Students and teachers' difficulty of seeing the generality embedded in a particular mathematical object has been noted in the literature (e.g., Yopp & Ely, 2016), leading some researchers to suggest that generic examples often have "opaque nature" (Rø & Arnesen, 2020),

Complementary to Mason and Pimm, we highlight the idea of imagining a particular in the general. The two counterexamples – Dina's construction and Olga's example of $n = 17k$ are general counterexamples because they provide a way to generate infinitely many specific counterexamples. The two general counterexamples require one to imagine or deduce how a particular case can be generated from a general description to come up with a specific counterexample. However, in some cases, the generality of these counterexamples appeared unconvincing just for the reason of *not* being particular. Thus, when seeking a counterexample, and being aware that "one is enough", we demonstrated cases where teachers were unable (or unwilling) to see the particular in the general.

6 On Transparency

Zazkis and Gadowsky (2001) used the notion of "transparency" in considering properties of numbers given as numerical expressions. They suggested that a representation can be *transparent* with respect to particular number properties, but *opaque* with respect to other properties. For example, the representation $a = (17 \times 20) + 3$ suggests that a leaves a remainder of 3 in division by 17, while $b = 7^3$ suggests that b is a cube. Yet, both representations are for the same number, $a = b = 343$. So, each representation is transparent with respect to some property but opaque with respect to another. It is also possible for an object to have no transparent representation to its properties. For example, while the representation $2k$ (for a natural number k) explicitly points to the evenness of the number, prime numbers do not have a representation that is transparent to their primality. This opaqueness was noted to present a difficulty for teachers (Zazkis & Liljedahl, 2004). In a similar vein, Arcavi (2003) reflected on the relative opaqueness of algebraic symbols, that while "sophisticated mathematicians may claim to 'see' through symbolic forms, regardless of their complexity", this is not the case for learners, who need to develop an ability of "reading of and through symbolic expressions." (p. 43).

Returning to the example of Selina (above), knowing that the number 437 is a product of 19 and 23 makes it transparent to a knowledgeable individual that it is a composite number without any testing. However, this feature was opaque to Selina, at least initially. Clearly, the notion of transparency and opaqueness is "in the eye of the beholder." A related mathematical knowledge is needed to recognize certain properties.

Consideration of transparency may explain why some counterexamples are easier accepted by teachers. A general counterexample of $17k$ is more transparent than that of 16. It shows *why* the expression $n^2 + n + 17$ is factorable, since substituting $17k$ in the expression, even mentally, immediately results in a multiple of 17, rejecting the possibility of the result being prime. Moreover, $17k$ shows a way to generate a whole class of (specific) counterexamples. A specific counterexample $n = 16$ is opaque since the fact that $16^2 + 16 + 17$ produces a composite number requires computation and then checking the result for primality. In addition, the counterexample of $n = 16$ does not point to possible existence of additional counterexamples.

Further, a construction of a general quadrilateral suggested by Dina is opaque to the type of resulting quadrilateral and its properties or the lack of thereof. Hence, even participating teachers who accepted the correctness of Dina's construction expressed a preference towards a sketch. These examples show that transparency of a counterexample can be related to generality as well as to particularity. That is, either general counterexample or a particular counterexample can be transparent (or opaque) to a beholder.

7 On Conviction and Epistemic Aims

Considering participants' individual epistemologies may provide an additional explanation of participants' different treatment of counterexamples. According to Chinn et al. (2011) individuals' epistemic cognitions are systems of beliefs about the nature of knowledge and processes of knowing. In psychology, knowledge is defined as "justified, true beliefs," that is "beliefs that accurately represent particular aspects of the world (at least approximately) and that are supported by adequate reasons" (p. 147). Knowledge, understanding and justification are central epistemic aims pursued by learners. *Understanding* involves grasping explanatory connections between elements of knowledge and information; *justification* relates to ways of using evidence to justify beliefs, and it is strongly tied to certainty. Rather than viewing certainty as a characteristic of knowledge itself, Chinn et al. (2011) define certainty as a "stance taken by an individual toward knowledge claims that are viewed as extremely well justified" (p. 156).

Importantly, Chinn and colleagues emphasize that while seeking certainty is one of the central epistemic aims, especially in mathematics and sciences, people may pursue other aims, such as avoiding false beliefs by withholding judgment altogether, judging claims as partially true, or adopting beliefs with a high ratio of truth to falsity. Consequently, individuals may seek justifications and explanations that adhere to their epistemic aims. These may include "explanations that explain the greatest range of data (even if there are some anomalous pieces of evidence that cannot be accounted for) or, instead, explanations that explain a narrower range of evidence but do so with few or no anomalies left unaccounted for" (Chinn et al., p. 155).

Returning to the participants' responses to the two statements and the proposed counterexamples, we note that although the statements dealt with school- level mathematical content, they were designed to be both unfamiliar and not immediately evident as true or false to the participants. We aimed at invoking uncertainty with respect to the truth-value of the statements, and to foster authentic evaluation of counterexamples.

Both counterexamples $n = 17k$ and $n = 16$ may provide an absolute conviction in the falsehood of the statement, but they address different epistemic aims. The counterexample $n = 16$ explains an extremely narrow range of evidence. While it may satisfy a modest epistemic goal of avoiding false beliefs, as evidenced in almost all participants accepting $n = 16$ as a legitimate counterexample, it does not allow for pursuing loftier epistemic aims, such as explanation or justification *why* the statement is false. This may explain the expressions of dissatisfaction with the specific counterexample by our participants.

The expression $n = 17k$ represents a family of counterexamples and a way to generate them, even though it does not account for all possible counterexamples. This general counterexample advances a stronger epistemic aim than merely avoiding false beliefs: it supports achieving stronger conviction in the falsity of the statement and it provides a strategy for addressing similar problems. Although mathematically, both $17k$ and 16 achieve the same aim of disproving the statement, these counterexamples satisfy different epistemic aims for the participants.

In fact, similarly to the case of $17k$, Dina's description does not account for all counterexamples. There is no necessity to require that all four segments resulting from intersection of the two congruent diagonals are of different length. Two of the four segments composing the diagonals can be of the same length, as long as they are parts of the same diagonal – such quadrilateral would disprove Statement-A. Requesting that all four segments on the diagonals are of different lengths, like Dina did, aims to invoke a mental image of a non-symmetrical, general figure, bestowing greater conviction in the existence of a counterexample.

What kind of epistemic aims can be associated with Dina's counterexample to Statement-A? Our data show that Dina's description of how to construct a general counterexample did not allow most of the participants to achieve the epistemic aim of forming justified true belief about the truth value of the statement. In other words, for most of the participants Dina's counterexample lacked the justificatory, explanatory power for conviction.

Among the sources of conviction described in philosophical literature are perception, which is grounded in the sensory experiences, and introspection, which is grounded in one's internal experiences; memory, reasoning, and testimony of others (Steup, 2005). Without an accompanying sketch, Dina's counterexample excludes perception as a source of conviction. Similarly, the participants could not rely on their memory as s source of knowledge about properties of quadrilaterals, since Dina's counterexample was a general quadrilateral without any specific properties, that would be readily evident to the participating teachers. To evaluate Dina's argument, the participants had to rely on introspection and reasoning. However, reasoning with figural, geometric objects may be quite difficult to perform introspectively. It is worth noting, that participants who sketched a quadrilateral according to Dina's description, were more likely to accept her counterexample as legitimate.

8 Conclusions

Recent studies have brought forward nuanced accounts of students' and teachers' conceptions of proof and counterexamples. For example, students who rely on empirical evidence for proving general statements are often aware of limitations of this approach (e.g., Bieda & Lepak, 2014); prospective teachers are aware that a proof and a counterexample cannot coexist (Stylianides & Al-Murani, 2010); and teachers' judgment of correctness of students' examples and counterexamples reflects pedagogical as well as mathematical considerations (Buchbinder, 2018).

Similarly, we sought to explain variance in teachers' treatment of counterexamples – openness towards some counterexamples and doubtfulness toward others – by examining various sources of conviction (or lack of it) with respect to counterexamples. Our initial assumption, which was based on the work of Peled and Zaslavsky (1997), that general counterexamples would be more welcomed, was not confirmed: we did not observe strong preference toward either general or specific counterexamples. To understand this phenomenon, we applied several theoretical lenses, each of which illuminated some aspect of the phenomenon, but neither provided a single satisfactory explanation. A transparency of a certain counterexample as convincing evidence disproving a false claim, seems to be "in the eye of the beholder." What makes some counterexamples transparent and convincing could not be easily established by using a single framework. Collectively, these theoretical lenses allowed attaining a more refined understanding of participants' conceptions of counterexamples and of their convincing power. Deepening understanding of the underlying reasons for what makes counterexamples convincing, or not convincing, through application of a variety of theoretical tools (Simon, 2009) is a valuable asset for researchers and educators alike. This paper, therefore, highlights the need for the field for developing more comprehensive theoretical tools that can explain a broad range of phenomena, and encompass a wide range of aspects related to the convincing power of counterexamples.

Acknowledgments

This study was supported by the National Science Foundation, Awards No. 1711163, 1941720. The opinions expressed herein are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Arcavi, A. (2005). Developing and using symbol sense in mathematics. *For the learning of mathematics*, *25*(2), 42-47.

Bieda, K. N., & Lepak, J. (2014). Are You Convinced? Middle-Grade Students' Evaluations of Mathematical Arguments. *School Science and Mathematics*, *114*(4), 166-177.

Biza, I. (2011). Students' evolving meaning about tangent line with the mediation of a dynamic geometry environment and an instructional example space. *Technology, Knowledge and Learning*, *16*(2), 125-151.

Brown, S. A. (2018). Are indirect proofs less convincing? A study of students' comparative assessments. *The Journal of Mathematical Behavior*, *49*, 1-23.

Buchbinder, O. (2018). Who is right? – What students' and prospective teachers' responses to scripted dialog reveal about their conceptions of proof. In R. Zazkis, P. Herbst (Eds.), *Scripting approaches in mathematics education* (pp. 89-113). Springer, Cham.

Buchbinder, O., & Zaslavsky, O. (2019). Strengths and inconsistencies in students' understanding of the roles of examples in proving. *Journal of Mathematical Behavior*, *53*, 129–147.

Chazan, D. (1993). High school geometry students' justification for their views of empirical evidence and mathematical proof. *Educational Studies in Mathematics*, *24*(4), 359–387.

Chazan, D. & Herbst, P. (2011). Challenges of particularity and generality in depicting and discussing teaching. *For the Learning of Mathematics, 31*(1), 9-13.

Chinn, C. A., Buckland, L. A., & Samarapungavan, A. (2011). Expanding the dimensions of epistemic cognition: Arguments from philosophy and psychology. *Educational Psychologist, 46*(3), 141-167.

Courant, R. & Robbins, H. (1996). *What Is Mathematics?: An Elementary Approach to Ideas and Methods, 2nd ed.* Oxford University Press, pp. 86-87.

Fischbein, E. (1993). The theory of figural concepts. *Educational Studies in Mathematics*, 24(2), 139–162.

Hanna, G. (2000). Proof and its classroom role: A survey. *Atas do Encontro de Investigação em Educação Matemática-IX EIEM*, 75-104.

Harel, G., & Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. In F. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning.* (pp. 805-842). NCTM. Reston, VA: Information Age Publishing, Inc.

Harel, G., & Weber, K. (2020). Deductive reasoning in mathematics education. *Encyclopedia of Mathematics Education*, 183-190.

Mason, J. (2006). What makes an example exemplary: Pedagogical and didactical issues in appreciating multiplicative structures. *Number theory in mathematics education: Perspectives and prospects*, 41-68.

Mason, J., & Pimm, D. (1984). Generic examples: Seeing the general in the particular. *Educational Studies in Mathematics, 15*(3), 277-289.

Peled, I., & Zaslavsky, O. (1997). Counter - Examples that (only) Prove and Counter-Examples that (also) Explain. *FOCUS on Learning Problems in Mathematics*, *19*(3), 49-61.

Rø, K., & Arnesen, K. K. (2020). The opaque nature of generic examples: The structure of student teachers' arguments in multiplicative reasoning. *Journal of Mathematical Behavior, 58*, 100755.

Simon, M. A. (2009). Research Commentary: Amidst Multiple Theories of Learning in Mathematics Education. *Journal for Research in Mathematics Education*, *40*(5), 477-490.

Steup, M. (2005). Epistemology. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from http://plato.stanford.edu/entries/epistemology/

Stylianides, A. J., & Al-Murani, T. (2010). Can a proof and a counterexample coexist? Students' conceptions about the relationship between proof and refutation. *Research in Mathematics Education*, *12*(1), 21-36.

Watson, A., & Mason, J. (2006). *Mathematics as a constructive activity: Learners generating examples*. Routledge.

Weber, K., & Mejia-Ramos, J. P. (2015). On relative and absolute conviction in mathematics. *For the Learning of Mathematics*, *35*(2), 15-21.

Yopp, D. A., & Ely, R. (2016). When does an argument use a generic example?. *Educational Studies in Mathematics*, *91*(1), 37-53.

Zaslavsky, O., & Shir, K. (2005). Students' conceptions of a mathematical definition. *Journal for Research in Mathematics Education, 36*(4), 317-346.

Zazkis, R., & Chernoff, E. J. (2008). What makes a counterexample exemplary?. *Educational Studies in Mathematics*, *68*(3), 195-208.

Zazkis, R., & Gadowsky, K. (2001). Attending to transparent features of opaque representations of natural numbers. *NCTM*, 41-52.

Zazkis, R., & Liljedahl, P. (2004). Understanding primes: The role of representation. *Journal for research in mathematics education*, *35*(3), 164-186.