



HAL
open science

Refutation feedback on student proofs beyond counter-examples

Alon Pinto, Jason Cooper

► **To cite this version:**

Alon Pinto, Jason Cooper. Refutation feedback on student proofs beyond counter-examples. Fourth conference of the International Network for Didactic Research in University Mathematics, Leibnitz Universität (Hanover), Oct 2022, Hannover, Germany. hal-04026703

HAL Id: hal-04026703

<https://hal.science/hal-04026703>

Submitted on 13 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Refutation feedback on student proofs beyond counter-examples

Alon Pinto and Jason Cooper

Weizmann Institute of Science, alon.pinto@weizmann.ac.il

Feedback on students' proofs is often intended to promote proof comprehension, yet formulating such feedback is a delicate task. In this study we investigate how refutation can be used for this purpose. We propose to extend Lakatos' notion of heuristic refutation to feedback that contains a refutation argument, possibly incomplete. Such feedback is heuristic in the sense that interpreting and utilizing it invites mathematical reasoning that can contribute to development of proof comprehension. Based on data from a Real-Analysis course, we show the importance of considering different kinds of refutation that are not based only on counter-examples, and demonstrate the nuances and subtleties of formulating feedback based on such refutation. Our findings suggest how professors can purposefully tailor feedback for particular didactic goals.

Keywords: Teachers' practices at the university level, novel approaches to teaching, professor feedback, heuristic refutation, real-analysis.

INTRODUCTION

Students' engagement with proof outside class is a central aspect of proof-oriented mathematics courses (Rasmussen et al., 2021). Students listen to their professors as they present proofs for the first time in lectures, and are expected to continue studying these proofs after class (Pinto & Karsenty, 2018; Weber, 2012). Weber (2012) observes that mathematics professors expect undergraduate students to do substantial work in order to understand a proof after it was presented in class, dedicating up to two hours to review proofs that are under ten lines long. In addition to reviewing the material that was presented in class, students are typically also expected to invest a substantial amount of time in writing their own proofs as part of their coursework, in order to promote further learning of proofs presented during the lectures (Rupnow et al., 2021). Feedback on students' proofs is widely recognized as having a key role in shaping and facilitating student learning between lectures (Moore, 2016; Pinto & Karsenty, 2018; Rasmussen et al., 2021). According to Moore (2016), professors' feedback on students' proofs is important for conveying norms and expectations, and for directing students' attention towards certain facets of the material, thus promoting proof comprehension.

While there has been a surge of empirical research of undergraduate mathematics teaching practices over the last decade, most attention has been given to what transpires inside classrooms, and research of professors' feedback on proofs that students submit as part of their coursework is fairly limited (Rupnow et al., 2021). Moore (2016) investigated professors' grading of students' proofs, and found that substantial variation in the scores assigned to similar proofs reflected to what extent flaws in the proofs indicated flawed comprehension. Moore's findings were corroborated and elaborated by Miller et al. (2018), who also highlighted the link between proof grading and students' apparent comprehension of the proof. Both Moore (2016) and Miller et

al. (2018) concluded that when grading students' proofs, instructors were assessing not only the correctness of the proof or whether it adheres to the norms of proof writing, but also (and mainly) students' proof comprehension.

There is preliminary evidence that, similarly, feedback on proof aims to promote proof comprehension (Byrne et al., 2018; Pinto & Karsenty, 2018, 2020). Evidently, professors often opt to leave the required revision – or even the flaw – implicit in their feedback, and instead highlight certain locations in the proof, ask eliciting questions or request elaborations (Moore, 2016; Byrne et al., 2018; Pinto & Karsenty, 2020). Such feedback may be viewed as an invitation for the student to engage with the flawed proof not only to correct it, but also to promote proof comprehension. However, Byrne et al. (2018) examined students' interpretations of six types of feedback and found that when instructors' feedback did not explicitly provide the required revision, students typically failed both to identify the flaw in their proof and to understand how they are expected to revise it. Conversely, Byrne et al. (2018) observed that when feedback was explicit regarding the desired revision, students often followed the prescription in the feedback without being able to explain what was flawed in their original proof or how the revision addressed flaws.

In this paper we focus on a type of feedback not examined by Byrne et al. (2018), which was used extensively and specifically for promoting proof comprehension in a course we examined in previous studies (Pinto & Karsenty, 2018, 2020). The professor in this course, whom we call Mike, opted to provide students with incomplete refutations of their proofs, arguing that such feedback often enables students to identify and fix the flaw in their proof almost on their own, while also promoting comprehension of the proof. We recognize in student engagement with this kind of feedback an interesting and potentially useful extension of the notion of heuristic refutation (de Villiers, 2010; Komatsu & Jones, 2021; Lakatos, 1976; Pinto & Cooper, under review). Furthermore, we observed that Mike used various kinds of refutation arguments, suggesting that heuristic refutation can be extended beyond the case of counter-examples considered in the literature. In this paper we introduce the notion of *heuristic refutation feedback* (HRF) and explore the following question: What kinds of refutation, beyond counter-example, can be formulated as HRF, and how?

HEURISTIC REFUTATION FEEDBACK

The intricate connection between proof and refutation has a long and respected history, originating from the works of Lakatos (1976) that highlighted how an interplay between proving and refuting can generate new mathematical knowledge (Komatsu & Jones, 2021). Drawing on works of Lakatos (1976) and de-Villiers (2010), Komatsu and Jones (2021) use the notion *heuristic refutation* for a mathematical activity that goes back and forth between conjecturing, attempting to prove, discovering counter-examples, and revising the conjecture, its proof, or even the definition of mathematical objects at stake. Komatsu and Jones (2021) stress that the term heuristic comes to emphasize the revision activity that stimulates growth of knowledge following the refutation, not the refutation itself. Counter-examples play a key role in the heuristic

refutation process by providing a trajectory for moving forward: the conjecture or the proof needs to be revised so as to neutralize the counter-example.

Both Komatsu and Jones and de-Villiers follow Lakatos in addressing a process of discovery in which a statement is not known to be true and may end up being revised if refuted. Lakatos (1976) called refutation of the statement *global* refutation, as opposed to *local* refutations, which challenge only one step in the proof or some aspect of the domain of validity of a statement. We wish to extend the notion of heuristic refutation to students' flawed proofs. Here, the global statement is known to be valid and will not be revised, yet we will claim that a particular kind of feedback on flawed proofs can invite students to engage in heuristic growth of knowledge when making sense of the feedback, reviewing their flawed proof with respect to the feedback, and attempting to revise their proof. To draw students into such heuristic activity, the feedback would need to unequivocally show that the proof is invalid, while leaving space for heuristic activity. Accordingly, we define *heuristic refutation feedback* (HRF) as any feedback on a flawed proof that contains a mathematical argument that indirectly implies that the proof is invalid. Here we extend the notion of global refutation to include cases where the refutation does not invalidate the statement, yet does reveal a structural failure in the proof, indicating that a local fix may not suffice, and that a different approach may be required. We emphasize that this definition extends Lakatos's and de Villiers' notion of *heuristic refutation*, which refers solely to engagement with counter-examples (local or global). While the literature recognizes different types of proof, which may have "diverse pedagogical properties and didactic functions in mathematics education" (Hanna & de-Villiers, 2008, p. 332), literature on refutation is generally restricted to counter-examples, and little is known about other kinds of refutation and how these may be utilized in teaching, particularly in the context of heuristic refutation.

Our definition of HRF extends the notion of heuristic refutation also in how the refutation argument may be formulated. By definition, HRF contains a (possibly incomplete) refutation argument. The reconstruction of an incomplete argument can be a challenging task that entails inference and invention of implicit connections between the feedback and the proof. As such, it can be seen as a case of abductive reasoning, and illustrated with Toulmin's (1958) model of arguments, as discussed by Komatsu and Jones (2021). In this model, an argument includes, among other things, a claim (C), datum (D) that supports the claim, and a warrant (W) that describes how the datum supports the claim. Abductive reasoning, as discussed by Komatsu and Jones (2021), is a process that seeks to explain a surprising observation (claim) through inference of hypotheses (data) and recognition or invention of warrants. Komatsu and Jones (2021) distinguish between three types of abductive reasoning, according to whether students need to complete the missing datum or provide a missing warrant. While in the classroom activities Komatsu and Jones considered the teacher has a key role in orchestrating student reasoning, in the context of HRF, the only way to facilitate heuristic refutation is through careful formulation of the feedback. This includes not

only selecting what to refute in the student's proof, and constructing an adequate refutation, but also deciding which elements of the refutation argument to present and how. Thus, we extend Komatsu and Jones's use of Toulmin model in two ways, by allowing each of the elements of the refutation argument, including the claim, to be stated implicitly, not only completely omitted. We exemplify and discuss these variations below.

METHODOLOGY

Data for this study were collected in a Real-Analysis course. The professor (Mike) is a mathematician who has been teaching for more than two decades and has taught this specific course several times. Prior research on Mike's goals and expectations with respect to this course (Pinto & Karsenty, 2018) revealed that he intended his feedback on students' proofs to be restricted as much as possible to refutations, stating that this kind of feedback affords opportunities for students to develop proof comprehension by looking for their own errors and for ways to correct them; develop the practice of testing their reasoning by trying to refute it; and develop a notion of validity of a proof that is absolute and independent of the professor's personal inclinations. Every week, Mike assigned a list of propositions to prove and examples to construct. Students submitted these proofs and examples electronically every few weeks. Mike did not grade students' submissions but provided written feedback. Seven students volunteered to participate in this study. The data corpus included a total of 57 submissions (5-12 submissions per student), and Mike's 2709 markings and comments.

Our first step in the analysis was to review all Mike's feedbacks to locate those that qualify as HRF and identify the refutation arguments therein. Often the feedback did not provide a fully argued refutation. In some cases, it was first necessary to identify what was being refuted, in particular in cases where the feedback refuted an implicit statement in the student's proof. When disagreements between the authors arose, they were discussed until agreement was achieved. Drawing on Toulmin's model of argumentation (1958) and on its application to heuristic refutation (Komatsu & Jones, 2021), we decomposed refutation arguments into three components (*Claim, Datum, Warrant*). As refutation argument typically coincided to some extent with arguments in the students' proof, we occasionally applied Toulmin's model to arguments within the students' proofs as well. Comparing the arguments in the feedback and in the proofs helped highlighting subtle and nuanced aspects of Mike's formulation of HRF. In this paper we focus on HRFs in which the datum was not a counter-example and show how Mike used different kinds of refutation while formulating HRF.

FINDINGS

At the beginning of the course, Mike defined the real numbers (\mathbb{R}) as an extension of the rational numbers (\mathbb{Q}) that includes non-repeating decimals and asked that students will prove at home that this definition implies that every non-empty bounded subset of \mathbb{R} has a least upper bound (Proposition 1). Mike then showed in class how Proposition

1 implies that \mathbb{R} is connected. One student, Alex, wrote a proof of Proposition 1 that relied implicitly on the connectedness of \mathbb{R} , and Mike provided the following HRF:

HRF1 Your argument is not only wrong, but it cannot be fixed, because you are not using any definition of real numbers, and hence whatever you write applies to \mathbb{Q} [the rational numbers], for which the whole statement is false.

In this feedback, Alex not only learns that the proof is flawed, but is invited to verify that the theorem does not apply to rational numbers, then retrace the line of reasoning in the proof, only with rational numbers instead of real numbers, recognize where the connectedness of \mathbb{R} is implicitly used, and realize that the proof of connectedness of \mathbb{R} relies on the proposition she is trying to prove. While we have not included Alex's proof, we note that refuting it with a counter-example would not be straight-forward, since the proposition is correct, and the proof does not contain a false statement. Mike's feedback does not specify the critical flaw in Alex's proof or the required revision, but in inviting Alex to revisit the definition of real numbers, it provides her with a route for recognizing the flaw on her own. We refer to feedback that derives a false statement by adapting the proof or a part of it as *refutation by false implication*.

We stress that HRF need not rely on *global* refutations, as can be seen for example in the case of the Extreme Value Theorem (EVT), which posits that if f is a real-valued continuous function on a closed bounded interval I then f attains both a maximum and a minimum in I . In the lecture, Mike emphasized that EVT is not as obvious as it may seem, noting that an analogous statement for rational-valued functions would not be true, even if restricted to polynomials. In the home assignment, students were asked to prove that there exists a polynomial $f(x)$ with rational coefficients that does not achieve an extremum when restricted to rational values in the interval $I=[0,1]$ (Proposition 2). All the students' proofs of Proposition 2 roughly followed the same proof scheme, which can be described as follows:

- Step 1. Define a cubic polynomial f with rational coefficients.
- Step 2. Identify the critical points of f and ascertain that one or both are irrational.
- Step 3. Conclude that as a rational-valued function, f does not attain both a maximum and a minimum in I .

Notably, Step 1 and Step 2 do not necessarily imply Step 3, since f may attain its (global) maximum or minimum in I at the rational endpoints of I . Moreover, in general, polynomials may have both rational and irrational critical points. Thus, Step 3 should be warranted by showing that the cubic polynomial f achieves its maximum or minimum in $[0,1]$ at the irrational critical points identified in Step 2. Most students left the warrant for step 3 implicit, as illustrated for example in Bailey's proof. Bailey defined $f(x) = \frac{7}{4}x^3 - 2x^2 + \frac{1}{6}x$, calculated the roots of f' , and stated:

Utilizing the quadratic formula, we see that the roots of the derivative of this function are [...] irrational. It is clear that the maximum and minimum occur at these irrational roots,

and thus the polynomial does not achieve its maximum or minimum value for $x \in \mathbb{Q}$ over the interval between 0 and 1.

Mike responded on Bailey's proof in the following way:

HRF2 $f(x)=x$ does not have roots of the derivative (even among real numbers!) but it does achieve its maximum and minimum values on $[0,1]$.

Notably, the polynomial Bailey proposed achieves its extrema in I at irrational points, and therefore has no maximum or minimum when restricted to $I \cap \mathbb{Q}$. Moreover, the proof does not contain an explicit false statement. Nonetheless, the implicit warrant may be incorrect. Mike's feedback retraces Bailey's line of reasoning, drawing on the same data – the derivative of $f(x)$ has no *irrational* roots – while replacing the polynomial f proposed by Bailey with $f(x)=x$, thus seemingly reaching a proof to an analogous statement, which is nonetheless false. Unlike HRF1, here the refutation is *local* as the proof could be readily fixed by adding an explicit (correct) warrant.

We note that formulating HRF to Bailey's proof entails attributing a flawed warrant to the unwarranted claim "*It is clear that the maximum and minimum occur at these irrational roots*". The warrant Mike's feedback attributes is: *For every polynomial f , if $f'(x)$ has no rational roots then $f(x)$ does not obtain a maximum or minimum in $I \cap \mathbb{Q}$.* We note that Mike could have chosen to attribute other warrants, for example: *For every closed bounded interval $[a,b]$, if $f'(x)$ has no roots in $[a,b]$ then $f(x)$ does not obtain a maximum in $[a,b] \cap \mathbb{Q}$.* In fact, this alternative path is reflected in Mike's HRF to Adrian's proof, which presented a line of argument very similar to Bailey's:

HRF3 These are critical points, but what makes you think that the maximum and minimum values of this f are achieved on $[0,1]$ at these points? The points do not even depend on the interval! Do you mean that the maximum and minimum values on every interval $[a,b]$ are the same? But this cannot be, because the polynomial is unbounded both above and below.

Attributing a false warrant to justify an unwarranted claim in a student's proof is not the only way to formulate HRF. In some cases, Mike's feedback altered data used explicitly but invalidly in the student's proof, as evident in his feedback to Dylan's proof of Proposition 2. Dylan, defined $f(x) = \frac{2}{3}x^3 + 2x^2 + x$ and stated:

To show [Proposition 2] we can demonstrate that the maximum value of the polynomial in this interval has no corresponding point in the specified domain. Since neither solution [of the equation $f'(x)=0$] is rational, we conclude that no least upper bound exists.

Mike responded to Dylan's proof in the following way:

HRF4 Note that both values of x [in which $f'(x)=0$] are outside the interval $[0,1]$. Thus, according to your logic, the range of your function does not have the least upper bound even over the real numbers. Contradiction?

Dylan's proof is different from Bailey and Adrian's proofs in that the polynomial it proposed is in fact a non-example, as the roots of f' reside outside the interval I , which

implies that f is monotonic on I and thus obtains both its minimum and its maximum at the (rational) endpoints of I . Rather than refuting Dylan's proof by a counterexample, Mike's feedback drew on the misused data – the critical points – highlighting that their irrationality plays no role in the proof, and therefore an analogous line of reasoning could be applied to the same data only with f as a real-valued function, and reach the same conclusion in contradiction to EVT.

The cases considered so far admitted a wealth of refutation arguments. But, in some cases, formulating refutation feedback was not straight forward. For example, Charlie's proof defined $f(x) = \frac{1}{3}x^3 - \frac{1}{2}x$, found a root x of $f'(x)$, and stated:

You can see that the first derivative equals 0 at x (x is between 0 and 1). The second derivative is positive at x , indicating that we've found a minimum in this interval, and x is not rational. However, the second derivative equals 0 at $x = 0$. In order to ensure that we have found a minimum for the interval $[0, 1]$, we still need to check that the value of $f(x)$ is less than $f(0)$.

Unlike the proofs of Alex, Bailey and Dylan, Charlie's proof provides an explicit warrant to why the irrationality of the critical point of f implies that f , as a rational function, does not obtain a minimum in the interval I . Charlie checks the sign of the second derivative of f at the irrational critical point $x \in I$, and rightly concludes that x is a local minimum of f . However, Charlie also notes that the second derivative of f is negative at every point of I , except that it vanishes at zero, and wrongly argues that in order to show that x is a minimum of f in I it is necessary and sufficient that $f(x) < f(0)$. Notably, whereas Charlie's line of reasoning is not valid, its conclusion for the particular f and I is true: x is indeed the unique minimum of f in I , and 0 is the unique maximum. Thus, showing that Charlie's reasoning is not always true entails the non-trivial task of constructing (or suggesting the existence of) an example in which all the data Charlie drew on can be used in the same way, leading to a false conclusion. Mike's feedback does just that:

HRF5 Apparently you see some connection between the sign of $f''(0)$ and extremal values. Here is a counterexample: Consider f on the closed interval $[0, 10]$. It has no local maxima, its 2nd derivative is positive on $(0, 10]$, and $f(0)=0$ is not a maximum, since, say $f(2) = 5/3 > 0$. Thus, according to your logic, the function does not achieve a maximum value on $[0, 10]$.

By expanding the interval from $[0, 1]$ to $[0, 10]$, Mike's feedback demonstrates that the maximum of f is not necessarily achieved at a point where the second derivative is non-negative. The warrant Mike's feedback attributes to Charlie's proof and refutes is: *The minimum of f in I is achieved at points in which the sign of f'' is not negative*. Notably, this warrant is alluded to at the beginning of Mike's feedback to Charlie: "Apparently you see some connection between the sign of $f''(0)$ and extremal values".

So far, we have examined HRF based on *global* and *local* refutations by false implication. Another type of refutation argument Mike used in his feedbacks is *refutation by contradiction*. This refutation argument, similarly to proof by contradiction, first assumed the student's proof or some statement therein is correct, only to reach a contradiction to the theorem that needs to be proved (and that is known to be true). For example, in one of the lectures, Mike presented the Peano curve as the limit of a sequence of curves $\phi_n: [0,1] \rightarrow [0,1]^2$. The curve ϕ_1 is defined as the curve that starts from the point (1,1), moves along the four edges of the unit square until it returns to (1,1), and then moves along the diagonal of the unit square to the point (0,0). The curve ϕ_{n+1} is defined recursively by replacing every diagonal line in (the image of) ϕ_n with a curve that consisted of 8 parts, as illustrated in Figure 2. Points on the unit interval mapped to horizontal or vertical segments on the Peano curve were labeled by 's'.

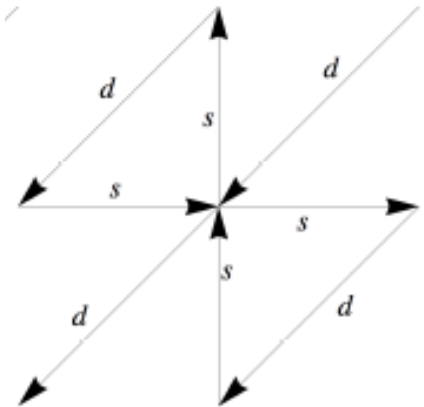


Figure 2: Construction of the Peano Curve

The students were asked to prove that the Peano curve is surjective. One student started the proof by claiming that “every point on the unit interval eventually falls into an interval labelled as *s*”. Mike responded with the following feedback:

HRF6 All ‘s’ points are mapped to points of the square with one coordinate rational; so, they don't cover the square. This contradicts the theorem.

Taking (implicitly) the student's claim as data, HRF6 posits that points of type ‘s’ are mapped to plane points with one rational coordinate. Combining these two pieces of data together leads to the false conclusion that the Peano curve is not surjective.

We conclude by pointing out that refutation by counter-example can be seen as a case of refutation by false implication. To illustrate this, we return to Proposition 1. One student stated in the proof that “*every closed set [of the real numbers] can be written as the union of only finitely many closed intervals*”. Mike addressed this statement:

HRF7 This is bluntly wrong: The Cantor set does not contain a single interval, but it is uncountable.

Here, the Cantor set is given as a counter-example to the false statement. At the same time, the argument can also be read as a false implication: applying the statement to the Cantor set implies that it is a finite union of closed intervals, which is absurd. We stress that the converse is not true, since, as demonstrated above, refutation by false implications can be used to refute proofs that do not admit false statements, and thus cannot be refuted by a counter-example.

DISCUSSION

This study, situated in the under-studied area of undergraduate mathematics teaching and learning outside class, aims to unpack nuances of a particular practice of

undergraduate mathematics teaching – providing written feedback on students’ flawed proofs. The work is guided by the premise that feedback on students’ flawed proof can support development of proof comprehension rather than merely support the writing of correct proofs. We have proposed heuristic refutation feedback (HRF) as an extension of heuristic refutation (de Villiers, 2010; Lakatos, 1976) to conceptualize the activity of interpreting and utilizing refutation feedback on flawed proofs, and have demonstrated that formulating such feedback can be a delicate and thoughtful practice.

We have extended the notion of heuristic refutation (de Villiers, 2010, Komatsu & Jones, 2021) in two ways. First, we go beyond the notion of abductive reasoning (Komatsu & Jones, 2021), where a claim based on observation was taken as given and it is up to students to propose datum and/or a warrant, and consider the heuristic activity of completing a refutation argument that may contain only a claim, only datum, only a warrant, or any combination thereof. Second, we go beyond refutation by counter-example to consider refutation by false implication. In this we are extending the space of pedagogical applications of refutation in mathematics education.

Formulation of HRF was shown to involve several different pedagogical decisions that relate to the construction and selection of the refutation argument. Typically, students’ flawed proofs can be refuted in more than one way. There is often more than one flaw in a flawed proof, and different flaws may indicate different issues of proof comprehension. Thus, formulating HRF may entail a decision about *what* to refute. Presumably, and this needs to be studied further, highlighting different flaws can provoke different engagement of students with their proofs, their flaws, their revisions and their comprehension thereof. Formulating HRF entails also decisions about *how* to refute. We have delineated two kinds of refutation arguments – *refutation by false implication* and *refutation by contradiction*, in addition to the familiar *refutation by counter-example*. We have demonstrated that in some cases, more than one kind of refutation argument is applicable. The literature suggests that different kinds of proof have different pedagogical advantages and afford different opportunities for learning (Hanna & de Villiers, 2008), and further research is need to investigate whether and in what sense this is also is also true for different kinds of refutation.

We have demonstrated how in some cases proofs may be flawed even if they do not contain an explicit invalid statement. Such proofs cannot be refuted directly by means of a counter-example, yet they may be refuted by identifying (or attributing) a flawed warrant and invalidating it. We have demonstrated how different warrants may be attributed to the same flawed proof, and lead to different HRF. We note that formulation of HRF entails also decisions about the extent to which different elements of the refutation argument (claim, datum, warrant) are explicated in the feedback (explicitly, implicitly or omitted). This aspect of the formulation of HRF is discussed in detail in a separate publication (Pinto & Cooper, under review). Thus far, the potential affordances of HRF for proof comprehension have only been substantiated theoretically (Pinto & Cooper, under review). Empirical research on how students

engage with this kind of feedback, and how this engagement can contribute to the development of proof comprehension remains for future research.

REFERENCES

- Byrne, M., Hanusch, S., Moore, R. C., & Fukawa-Connelly, T. (2018). Student interpretations of written comments on graded proofs. *International Journal of Research in Undergraduate Mathematics Education*, 4(2), 228-253.
- de Villiers, M. (2010). Experimentation and proof in mathematics. In G. Hanna, H. N. Jahnke, & H. Pulte (Eds.), *Explanation and proof in mathematics* (pp. 205-221). Springer.
- Hanna, G., & de Villiers, M. (2008). ICMI Study 19: Proof and proving in mathematics education. *ZDM – Mathematics Education*, 40(2), 329-336.
- Komatsu, K., & Jones, K. (2021). Generating mathematical knowledge in the classroom through proof, refutation, and abductive reasoning. *Educational Studies in Mathematics*, 109(3), 567–591.
- Lakatos, I. (1976). *Proofs and refutations*. Cambridge University Press.
- Miller, D., Infante, N., & Weber, K. (2018). How mathematicians assign points to student proofs. *The Journal of Mathematical Behavior*, 49, 24-34.
- Moore, R. C. (2016). Mathematics professors' evaluation of students' proofs: A complex teaching practice. *International Journal of Research in Undergraduate Mathematics Education*, 2(2), 246-278.
- Pinto, A., & Cooper, J. (under review). "This cannot be" – Refutation feedback and its potential affordances for proof comprehension.
- Pinto, A., & Karsenty, R. (2018). From course design to presentations of proofs: How mathematics professors attend to student independent proof reading. *The Journal of Mathematical Behavior*, 49, 129-144.
- Pinto, A., & Karsenty, R. (2020). Norms of Proof in Different Pedagogical Contexts. *For the Learning of Mathematics*, 40(1), 22-27.
- Rasmussen, C., Fredriksen, H., Howard, E., Pepin, B., & Rämö, J. (2021). Students' in-class and out-of-class mathematical practices. In *Research and Development in University Mathematics Education* (pp. 216-237). Routledge.
- Rupnow, R., Hegg, M., Fukawa-Connelly, T., Johnson, E., & Weber, K. (2021). How mathematicians assign homework problems in abstract algebra courses. *The Journal of Mathematical Behavior*, 64, 100914.
- Toulmin, S.E. (1958). *The uses of argument*. Cambridge University Press.
- Weber, K. (2012). Mathematicians' perspectives on their pedagogical practice with respect to proof. *International Journal of Mathematical Education in Science and Technology*, 43(4), 463-482.