Obligations and Norms: How Instructors Respond to Students' Correct But Non-Traditional Proofs

Andrew Kercher
Simon Fraser University

Anna Marie Bergman
Fort Lewis College

Rina Zazkis
Simon Fraser University

Proof is a central pillar of mathematical practice, and so teaching proof to undergraduate students is a necessary responsibility for undergraduate mathematics instructors. Our study contributes to research on how instructors can achieve this goal by examining their responses to student-generated existence proofs. In particular, the simple logical structure of an existence proof allows for proof-writing norms to be the focus of attention. By analyzing instructors' responses to students' existence proofs in terms of norms, values, and professional obligations, our study attends to the ways in which norms are (or are not) supported by instructors in undergraduate mathematics classrooms.

Keywords: Proofs, Existence Proofs, Professional Obligations, Norms, Values

Among mathematicians, it is not uncommon to agree with Ziegler's (2013) observation that "proofs are the heart and mind of mathematics. They are the pillars upon which the structure of mathematics rests" (p. 130). Among students entering undergraduate mathematics programs, however, the centrality of proof may not be so self-evident; given the scarcity of proving activity featured at the secondary level, this limited perspective is perhaps understandable. It becomes the work of undergraduate mathematics instructors, then, to foster a fuller perspective on proof that more closely aligns with the actual work of mathematics—to illustrate to their students the nature and importance of proof by teaching them what proofs are, how to write them, and what role they serve.

This is not a simple task. Countless studies have explored the difficulties undergraduate students navigate as they come to understand proofs as a concept (e.g., Dawkins & Weber, 2017; Stylianides & Stylianides, 2009; Selden & Selden, 2013). Other researchers have spotlighted particular types of proofs, such as proof by contradiction (e.g., Chamberlain & Vidakovic, 2021), contraposition (e.g., Antonini & Mariotti, 2008) or mathematical induction (e.g., Norton et al., 2022), that are notoriously challenging for students.

By contrast, proving a statement of existence is perceived to be an approachable process for mathematical novices; such proofs are direct, typically constructive, and require only the provision of a single example (Buchbinder & Zaslavsky, 2019). Perhaps for this reason, there is comparatively little research on how undergraduate mathematics students develop an understanding of this type of proof.

But the simplicity of an existence proof is beneficial for researchers interested in how undergraduate students understand the form and function of proofs in mathematics. Because they are not encumbered by logical constructions that are difficult or unfamiliar to them, these students are free to produce existence proofs that adhere most closely to what they perceive a proof "should be like". In this way, existence proofs provide a clearer lens into the norms and values around proof-writing that students have begun to develop.

Simultaneously, the way that undergraduate mathematics instructors read and respond to students' proofs in the classroom—especially those that do not adhere exactly to conventional proof structure—can reveal the mathematical norms and values that inform their teaching. But undergraduate mathematics instructors also have responsibilities towards their institution or their students' learning that might affect how they respond to proofs. Our study examines instructors' reactions to non-traditional student-generated existence proofs to better understand what aspects of proving those instructors value and thus promote in their classrooms. This answers the research question: *What do mathematics instructors consider when evaluating student-written existence proofs and how do their considerations relate to normative conventions for proving?*

Background

To answer our research question, we must first consider what values and norms mathematicians might conventionally hold. Dawkins and Weber (2017) conducted a thorough review of literature in order to present such a list of values, and the associated norms, with respect to proving. The four values they identified were:

1. Mathematical knowledge is justified by a priori arguments.
2. Mathematical knowledge and justifications should be a-contextual and specifically be independent of time and author.
3. Mathematicians desire to increase their understanding of mathematics.
4. Mathematicians desire a set of consistent proof standards.

Each of these values begets a collection of relevant norms, that is, suggested practices for fulfilling the associated value when engaged in proving. For example, the norm "*Irrelevant statements are not presented in the proof*" is a convention used by mathematicians to support the value for increasing understanding; Dawkins and Weber argue that "adding irrelevant statements or assumptions will confuse the reader as she struggles to find out how this is relevant" (p. 131).

Rupnow and Randazzo (2023) extend the work of Dawkins and Weber (2017) by interviewing mathematicians about the values and norms they hold with respect to defining as opposed to proving. They propose an additional value: that mathematicians desire clarity in and for communication when writing definitions.

But when mathematics teachers must respond to student-written proofs or definitions, they are often at odds—they must introduce their students to conventional mathematical values and norms while still attending to the needs of their students, their classroom, and their institution. To understand how mathematics instructors might balance these conflicting needs, we turn to Erickson et al.'s (2021) concept of *professional obligations*. The authors explain: "If norms describe expectations for action held in common by members of a profession, professional obligations help us understand why actions at a particular point and time may deviate from those expectations" (p. 192). Mathematics teachers might experience obligations to the *individual student*, the *social classroom*, the *institution*, or to the *discipline of mathematics*. For example, a teacher might endorse a non-normative proof out of obligation to an individual student if they perceive that doing so would have a positive effect on that student's emotional or intellectual needs. Thus, a mathematics teachers' decisions in the classroom can be seen as aligning with a mathematical norm (and thus, promoting a particular mathematical value), or else as fulfilling a professional obligation to some other entity.

Methodology

Participants and Data Collection

Our data comes from two sources: an online survey and semi-structured interviews, both with mathematics instructors. The survey collected participants' ($n = 86$) reactions to five different student responses; these responses were proofs of the statement, "There exist four different prime numbers the sum of which is prime," and are provided in Table 1. The presented responses were inspired by student work submitted in a course intended to serve as an introduction to proofs.

Table 1. The five student responses assessed by participants in the study.

Participant	Proof
Alpha	True. In order to prove this exists we only have to show one valid example of the statement. *a, b, c, d, e* in "primes" $a + b + c + d = e$ Example: 2 + 5 + 7 + 17 = 31 2, 5, 7, 17, and 31 are all primes. Therefore, there does exist a case where four distinct/ different primes sum to equal a prime number.
Beta	True. For example, consider the sum of prime numbers 2 + 3 + 7 + 11 = 23 which proves that there exists four different prime numbers the sum of which is prime. But there also exists four different prime numbers the sum of which is not prime. For example, 2 + 5 + 7 + 11 = 25.
Gamma	True. 2, 3, 5, and 7 are all different positive primes. Their sum, 2 + 3 + 5 + 7 = 17, is also a prime number. This example proves the initial statement that there exists a (at least one) solution. Note that 2 must be one of the primes in the sum: 2 is the only even prime number, so if none of the four primes is 2, then we have odd + odd + odd + odd = (2a + 1) + (2b +1) + (2c + 1) + (2d + 1) = 2(a + b + c + d + 2) = 2k where k is an integer, hence the result is even and cannot be prime (unless k = 2, except all odd primes are greater than 2, so the sum will be greater than 2 making this not possible).
Delta	True, 2 + 3 + 5 + 7 = 17, all primes.
Epsilon	This statement is true. Proving by contradiction, let us assume that this statement is false, then it becomes, "There don't exist four different prime (positive) numbers the sum of which is prime." We can disprove this with a single example, such as 2, 5, 7, and 17, which sum to 31, a prime number. As the inverse statement is false, we can see that the original statement, that there exists at least one group of prime numbers that sum to a prime, is true.

We note that, despite one minor (unintentionally included) error in Gamma's proof (the phrase "unless $k = 2$" should be "unless $k = 1$"), each proof is mathematically correct and logically sufficient to prove the existence of the requisite set of prime numbers. The proofs were chosen for the survey because they did not adhere to the previously observed norm for minimality found in existing research literature (or, in Delta's case, pushed the limits of what a lower bound on minimality might be).

Participants were prompted to "Imagine that you were teaching a math class, and you received the following five homework submissions as an argument or justification for whether the statement is true or false". Under these circumstances, participants were asked to rank the five student responses from "best to worst", under whatever criteria they deemed to be appropriate. More importantly, they were also asked to explain the rationality for their ranking.

The interviewees ($n = 5$) represent primarily a convenience sample, but one that targeted a variety of backgrounds both in learning and teaching mathematics. Interviewees completed the same survey as described above, but in the presence of a member of the research team. Doing so allowed the interviewees to explain in more detail their method for ranking the student responses; it also gave the research team member conducting the interview the opportunity to ask clarifying questions and prompt the interviewee to provide additional insight. Each of the interviews was audio-recorded and transcribed.

Data Analysis

Because the student responses were chosen for their relationship to the norm for minimality in mathematics, participants' explanation of their rankings were first coded by one member of the research team according to whether or not they indicated approval of the amount of content included in each student response. When participants endorsed a student's proof that was not minimal, their rationale was associated with a corresponding professional obligation. In other situations (e.g., when a non-minimal proof was rejected, or when Delta's proof was rejected despite its minimality) techniques from thematic analysis (Braun & Clarke, 2019) were used to identify commonalities between participants' responses. Through constant comparison, these commonalities developed into themes that sometimes aligned with existing norms or values that could be used to add context to the participants' reasoning. Next, a second member of the research team coded the participant explanations following the same procedure and the two researchers met to resolve any differences. An identical process was used to code and recode the interview transcripts.

Findings

In this report, we focus on three of the five responses that provide the most fertile analysis: Gamma, Delta, and Epsilon. We present participants' responses to each of the three existence proofs in separate subsections. Then, we provide a synthesized analysis of these responses.

Gamma

Gamma's response was ranked as the best by a majority of participants ($n = 63$) and by every interviewee. This overwhelming support of Gamma's proof, despite the fact that it is not normatively minimal, can be attributed to several different factors.

First, participants often identified Gamma as having produced the "best mathematical response" by "investigating and extending the problem." In particular, Gamma's observation that one of the prime numbers in the sum must be 2 was more than just an arbitrary mathematical fact appended to the end of the actual proof; participants considered this contribution a "corollary" that unpacks the "underlying structure of the statement rather than just providing an example." Participants stressed that Gamma's extra result was mathematically interesting and alluded to a more general explanation for how a set of primes meeting the requirements of the theorem could be found. This perspective was exemplified in Alan's interview when he said, "I don't particularly like the question. It doesn't do anything. It's nothing. There's no theory behind it. And Gamma, as I said, is the closest to coming up with a theory that makes the question interesting." Alan appears to accept Gamma's breach of the minimality norm because he feels an obligation to the discipline of mathematics, in the sense that it is a disservice to the subject to spend time proving inconsequential results. To Alan, mathematical proof is substantive and worthwhile if it is built on or contributes to existing theory; in his opinion, Gamma's corollary does exactly this.

Other participants ranked Gamma's proof as the best because they ordered the student responses according to the amount of mathematical understanding they revealed. These responses demonstrate an obligation to the student: unnecessarily longer proofs can be allowed when they provide an instructor with insight into their students' ways of thinking. Some participants reported that Gamma's extra work showed "the most conceptual understanding behind the answer"—that is, it revealed the extent of Gamma's mathematical subject matter competency. Other participants noted that Gamma probably "understands that they were done after the first sentence, but they enjoyed proving a little extra." In this sense, Gamma isn't displaying mathematical content knowledge but rather a "meta-theoretical understanding of the role of an example." Gamma's higher-level grasp of the role and methods of proof is evident, somewhat ironically, only because of Gamma's explicit violation of the norm for minimality in proving.

Finally, a number of participants recognized Gamma's proof as the best response because they quantified "best" according to how useful the student's proof would be as a pedagogical tool for promoting mathematical discourse in a classroom discussion. One participant explained that "Gamma […] immediately provided an example in the clearest way possible. The extra work they did actually put them above Delta because it gives me, as an instructor, ways to extend the question in possible follow-up discussion." This sentiment was echoed by Anthony in his interview: "So, in that sense, Delta and Gamma's responses are both good and I think it would be interesting to talk about which one would be preferred […] as a class." Anthony ranked both Delta and Gamma highly out of obligation to the social classroom. That is, he foresaw a chance for his students to collaboratively negotiate a sociomathematical standard for "how much" information should be included in a proof. The length of Gamma's response was acceptable exactly because it was not minimal, and thus could be juxtaposed against Delta's particularly extreme minimalism.

In fact, Gamma's response was never ranked as the worst by any participant and was only identified as the fourth best response a total of eight times. Most often, participants who ranked Gamma especially low took issue with his mathematical mistake (writing "unless $k = 2$" instead of "unless $k = 1$") rather than the length of their response. However, one participant linked the two: "The most concise answers are the best. [Answers like Gamma's], while to be encouraged since they are exploring the problem further, are more likely to lead to responses including incorrect statements."

Delta

Delta's response was ranked as the best by five participants. Their rationale for doing so was consistent: these participants valued Delta's "directness and simplicity," and variously described his response as "concise," "succinct and clear," and "short and to the point". This demonstrates a very clear preference for proofs that adhere as strictly as possible to the norm for minimality. Alan provided another perspective on Delta's work; although he also admitted that he liked it "for its brevity," he explained that brevity was only warranted because "this question is not significant in any way, so I think it's—the response given is on par for the level of the question."

On the other hand, Delta's response was ranked as the worst by half of the total participants ($n = 43$). These participants took a much more negative view of Delta's level of explanation, describing it as "incomplete" or "insufficient". Importantly, this appeared to be a mathematical shortcoming—one participant felt that Delta "provided no warrants at all for how this [example] justifies the claim." Another explained that, although "Delta's example is technically correct," he should "include more structure to make this a more formal argument." These responses indicate that participants were not necessarily concerned with Delta's understanding but instead with the understanding of those who might need to read and interpret his proof. This sentiment was outlined explicitly by Kenneth during his interview:

> For me, proof in mathematics is about communication of ideas. And for me, I—if I didn't have the question in front of me, I didn't know what we were trying to prove, I would have *no* idea what we were trying to do here, right?

This aligns with earlier results from Rupnow and Randazzo (2023) that mathematicians value communication but applied to proofs rather than definitions. However, it can also be interpreted as a type of obligation to the discipline of mathematics. Kenneth appeared to believe that proofs are not just a means of establishing deductive truths, but also a means of sharing those truths with fellow mathematicians. To do this, proofs need to be self-contained; this perspective is, in fact, one of the important mathematical values delineated by Dawkins and Weber (2017). From this perspective, Kenneth's reaction to Delta's proof is an example of competing mathematical norms: A proof must be minimal, but not so minimal that it cannot stand alone as an independent mathematical object.

Other participants who ranked Delta's response last were more concerned with Delta's personal understanding; in the most extreme case, one participant wondered if Delta had even copied their proof from a peer. These participants explained that Delta's work "gave very little insight into their thinking," and thus did not allow them to "assess what they understand." Kenneth's interview added a caveat to this assessment of Delta's understanding. Kenneth recognized that he "might accept this from a graduate student because I would assume that they have a certain baseline level of proof knowledge, but from an intro to proof course, my expectations are a little more, I guess." This illustrates the importance of context for identifying mathematical norms, especially in mathematics classrooms.

Epsilon

Epsilon's response was ranked as the best in eight participants' responses. Interestingly, despite the fact that Epsilon was ranked first in so few responses, there were still two clearly distinct schools of thought as to why it should be placed above its peers.

First, some participants thought that Epsilon's proof was simply the least incorrect. These responses typically took issue with Gamma's mathematical error and Delta's insufficient explanation, leaving Epsilon as the only "valid proof with no errors." Kenneth added some context to this perspective when he explained that

> Epsilon is clearly communicating everything that they're doing. So they—they're clear about what method of proof they're going to use. […] I don't love that they chose proof by contradiction, but they've done it in a nice way and it's clear to me what they're doing.

That is, Epsilon's proof may not be stylistically preferable but is still technically correct. Anthony also summarized this somewhat conflicted appraisal of Epsilon's proof by admitting that "if I was going to write a contradiction proof for this existence theorem, which is bizarre, it would probably look like Epsilon's. I just don't think that's a good idea."

On the other hand, some participants ranked Epsilon's response as the best specifically because of its logical complexity. One participant characterized Epsilon's proof as the "most sophisticated," while another described it as the "most formal." Importantly, this sophistication was seen as evidence of Epsilon's mathematical understanding: not only did he clearly understand how a proof by contradiction should be logically structured, but he was also able to use a proof by contradiction in a surprising and unexpected way. Alan provided a counterpoint to this interpretation of Epsilon's understanding when he observed that "it seems like Epsilon is, like, over formalizing. Maybe because they expect, you know, this is a math class and I'm supposed to be really formal." Then, the use of proof by contradiction could be seen as an explicit lack of understanding. In either case, however, the instructors are acting out of obligation to the student; Epsilon's clearly non-minimal proof is important to allow because it ultimately provides evidence that further conversation should take place with Epsilon.

Epsilon's response was the worst-rated proof in almost as many cases as Delta's ($n = 36$). Almost universal criticism was directed at Epsilon's choice of proof by contradiction, which participants found "labored," "unnecessarily complex," or simply "less clear" than other choices. This was a mathematical consideration; these participants were not concerned with why Epsilon felt that a proof by contradiction was warranted but were simply dissatisfied with the resultant mathematical object. Other participants, however, did identify Epsilon's choice of proof framework as insight into his (lack of) understanding. For example, one participant wondered whether Epsilon might be "uncomfortable with existence proofs". Another inferred that Epsilon was "proving things a bit ritualistically—they seem to be copying a known proof structure without considering whether it is the best way to prove the theorem in question". This was the basis of Gavin's poor appraisal of Epsilon's proof. During his interview, Gavin explained that Epsilon "complicated the solution unnecessarily. I don't think they understand clearly the process of solving—or sorry, proving this question."

Discussion

Our research question asked: *How do mathematics instructors respond to student-written existence proofs and how do their responses relate to normative conventions for proving?* Very commonly, non-minimal existence proofs that included technically extraneous information were accepted by participants. This is best illustrated by the overwhelming approval of Gamma's proof. Participants appreciated that Gamma's additional contribution was mathematically interesting, revealed a deeper understanding, and could serve as a foundation for classroom discussion. The latter two points are unique to the pedagogical context in which Gamma's work is embedded, but the former point is strictly a mathematical consideration. But minimality is also a mathematical consideration, and so in this way, reactions to Gamma's proof illustrate how two mathematical norms for proving can be in conflict even when they are both in service of the same value: that proofs should increase a mathematicians' understanding. A similar observation was made about reactions to Delta's proof, in which minimality was at odds with the need for the proof to make sense as an independent object.

Despite Gamma's positive reception, other non-minimal proofs (i.e., Epsilon's) were largely rejected by the same participants. This is in spite of the fact that, as pointed out by some interviewees, Epsilon's proof does in fact meet the standard for minimality (in terms of a proof by contradiction) and is revealing of Epsilon's understanding. Ostensibly, Epsilon's proof also could be used to stimulate a classroom discussion. We hypothesize that the difference between reactions to Gamma and Epsilon's proofs lies in the fact that Epsilon's proof added logical complexity without any accompanying mathematical insight.

Conclusion

Ultimately, participants in this study illustrated that mathematics instructors are often required to attend to non-mathematical obligations when considering student-written proofs. One limitation of this study was that much of the pedagogical context that might have led to these obligations was not provided in the survey—some participants commented that it was difficult for them to appraise the student responses without more information on the structure and purpose of the class, for example. Future research might include fewer student responses, but within a more elaborate fictional setting to allow instructors the opportunity to make more specific judgements.

References

Antonini, S., & Mariotti, M. A. (2008). Indirect proof: what is specific to this way of proving?. *ZDM*, *40*(3), 401-412. https://doi.org/10.1007/s11858-008-0091-2

Braun, V., Clarke, V., Terry, G., & Hayfield N. (2019). Thematic analysis. In Liamputtong, P. (Ed.), *Handbook of research methods in health and social sciences* (pp. 843-860). Springer. https://doi.org/10.1007/978-981-10-5251-4_103

Buchbinder, O., & Zaslavsky, O. (2019). Strengths and inconsistencies in students' understanding of the roles of examples in proving. *The Journal of Mathematical Behavior*, *53*, 129-147. https://doi.org/10.1016/j.jmathb.2018.06.010

Chamberlain Jr, D., & Vidakovic, D. (2021). Cognitive trajectory of proof by contradiction for transition-to-proof students. *The Journal of Mathematical Behavior*, *62*, 100849. https://doi.org/10.1016/j.jmathb.2021.100849

Dawkins, P. C., & Weber, K. (2017). Values and norms of proof for mathematicians and students. *Educational Studies in Mathematics*, *95*, 123-142. https://doi.org/10.1007/s10649-016-9740-5

Norton, A., Arnold, R., Kokushkin, V., & Tiraphatna, M. (2022). Addressing the cognitive gap in mathematical induction. *International Journal of Research in Undergraduate Mathematics Education*, 1-27. https://doi.org/10.1007/s40753-022-00163-2

Rupnow, R., & Randazzo, B. (2023). Norms of mathematical definitions: imposing constraints, permitting choice, or both?. *Educational Studies in Mathematics*, 1-18.

Selden, A., & Selden, J. (2013). The genre of proof. In M. Fried & T. Dreyfus (Eds.), *Mathematicians and mathematics educators: Searching for common ground* (pp. 237–260). Springer.

Stylianides, G. J., & Stylianides, A. J. (2009). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education*, *40*(3), 314-352. https://doi.org/10.5951/jresematheduc.40.3.0314

Ziegler, G. M. (2013). *Do I count?: Stories from mathematics*. CRC Press. https://doi.org/10.1201/b15150