The Use of Proof Comprehension as a Self-Assessment Tool in Undergraduate Proof Writing

Mitchell Alvarado
CSU, Bakersfield

Aubrey Kemp
CSU, Bakersfield

Research has shown that self-assessment techniques can improve understanding and academic performance for mathematics students. As proof construction, validation, and comprehension have proven to be facets of mathematics with which students can use more learning opportunities, Mejía-Ramos et al. (2012) present an assessment model for proof comprehension. For our project, this framework is utilized in understanding proof self-assessment, with the goal of aiding students in the process of proof construction and validation. In this preliminary report, we provide results from the first phase of our project, in which we surveyed seven university mathematics faculty about their approaches to proof construction and self-assessment. Results from this study are being used to generate a proof self-assessment tool through the lens of proof comprehension that undergraduate students may use to improve their proof writing strategies. Pedagogical suggestions are also provided based on these results.

Keywords: Self-Assessment, Proof Writing, Proof Comprehension, Evaluation, Criteria

Introduction

Recently, there have been many research studies conducted on undergraduate students' understanding of proofs, including the construction and writing of proofs (e.g., Azrou, 2015; Chamberlain & Vidakovic, 2021; Lew & Mejía-Ramos, 2019; Mejía-Ramos & Weber, 2019), validating proofs (e.g., Ko & Knuth, 2013; Powers et al., 2010), proof comprehension (e.g., Chamberlain, 2017; Dawkins & Zazkis, 2021; Selden & Selden, 2015; Weber & Tanswell, 2022; Yang & Lin, 2008) and assessment of proof comprehension (e.g., Mejía-Ramos et al., 2012; Cooley et al., 2024). Melhuish et al. (2022) and Stylianides et al. (2017) provide comprehensive reviews of research in many areas of proof in undergraduate mathematics education. Yang and Lin (2008) express it is important while constructing a proof to "have diverse ideas come to mind at the 'right time,' reorganize these ideas, and formally write the proof," (p. 73). The reorganization of these ideas is a cycle of writing, validating, reading and comprehending, and evaluating one's own proof. Selden and Selden (2015, 2017) discuss the interaction and comparisons of these various aspects of proof writing in undergraduate mathematics and point out there has been little research on how proof comprehension, construction, validation, and evaluation are related. As such, in this study, we seek to investigate how proof comprehension strategies may be used as a self-assessment tool in proof construction and evaluation.

Research shows students' academic performance can benefit from self-assessment and self-monitoring techniques (Andrade & Valtcheva, 2009; Beumann & Wegner, 2018; Kemp et al., 2020; Poehner, 2012; Stallings & Tascione, 1996) and students' proof writing can benefit from proof validation tasks (Powers et al., 2010) and self-explanation practice (Hodds et al., 2014). Further, students who use self-monitoring techniques have higher self-efficacy and self-assessment accuracy (Andrade & Valtcheva, 2009; Poehner, 2012; Stallings & Tascione, 1996). However, in some cases where students feel as though a grade is emphasized in the self-assessment, students may evaluate their work based on effort or focus on what may give them partial credit on an assignment (Kemp et al., 2020; Stalling & Tascione, 1996). Notably, in Kemp et al., (2020), the researchers describe a shift in how the participant uses her own criteria as she is evaluating her responses versus a peer's responses to proof comprehension questions.

These authors suggest future research can investigate how proof comprehension assessments for students influence their criteria in self-evaluation and understanding of proof writing. Beumann and Wegner (2018) investigates self-assessment in functional and real analysis on homework related to proofs. The authors present evidence that students can benefit from self-assessment tasks depending on several factors, and they note positive outcomes for self-assessment in small group tasks where the group can achieve full credit by hard work. They call for more research on and sophisticated implementations of self-assessment in higher mathematics education.

Beumann and Wegner (2018) advocate for self-assessment in that it should allow students to complete meta-tasks that are not marked by an instructor and help generate appropriate beliefs about mathematics, as self-assessment reinforces that checking the validity of a proof requires careful work and may depend on personal taste (i.e., the strategy used or amount of details given). Research shows students are better able to self-assess when given criteria on which to base their performance (Andrade & Valtcheva, 2009; Poehner, 2012), but Weber (2012) finds that mathematicians can greatly disagree on the validity of even simple proofs. When comparing mathematicians' and undergraduate students' beliefs about mathematical proofs, Lew and Mejía-Ramos (2019) report the following trends: (a) mathematicians believe proof language should obey conventions of academic language while students often either do not know these conventions apply to proof writing or are unaware they exist at all; (b) students face obstacles in understanding how mathematicians introduce objects in proofs; (c) students may not be aware of how important it is to focus on the context of a proof to decide how formal it should be. Since the criteria being used to evaluate a student proof can vary depending on who is evaluating the proof (e.g., an instructor, a peer) and what the purpose of the evaluation is (e.g., for effort, an assessment, peer-evaluation), it makes sense that students may not be aware of or understand the criteria under which they are being evaluated or may face obstacles in self-assessing and monitoring their progress (Stallings & Tascione, 1996; Poehner, 2012).

Based on the researchers' experiences and results from relevant research, our overall project seeks to investigate in what ways self-assessment through the lens of proof comprehension can improve proof construction and evaluation skills of undergraduate students. Specifically, we proceed with the assumption that self-assessment in proof writing necessarily requires proof comprehension skills (as one is reading their own proof), proof evaluation skills (as they are deciding if what they have written is logical and "complete"), and an awareness of what criteria, if any, is being used to evaluate their proof. This report provides preliminary results on the following research questions after a first phase of data collection investigating the perspectives of mathematics faculty on proof self-assessment: (a) What strategies and/or criteria do mathematicians use to self-assess proofs as they write them? (b) In what ways can proof comprehension assessment strategies be used in undergraduate self-assessment of proof writing?

Theoretical Framework

We focus in this preliminary report on the proof comprehension and assessment models presented in Yang and Lin (2008) and Mejía-Ramos et al. (2012). Yang and Lin (2008) provide a model of proof comprehension for proofs in high school geometry. The researchers use mathematicians' and mathematics teachers' perspectives to develop facets of reading comprehension in geometry proof. The authors emphasize that when students learn mathematical proof, there is often a gap between conjecturing and formally verifying, and their model aims to overcome this gap or eliminate limitations in proof writing. Mejía-Ramos et al. (2012) presents an assessment model for proof comprehension in undergraduate mathematics, which builds off of Yang and Lin's (2008) and Conradie and Frith's (2000) models, as well as trends found from interviews with mathematicians and other mathematics education literature. They provide a multidimensional model with seven types of proof comprehension assessment, divided by *local* and *holistic* types of assessment. The local types are *meaning of terms and statements*, *logical status of statements and proof framework*, and *justification of claims.* The holistic types are *summarizing via high-level ideas, identifying the modular structure*, *transferring the general ideas or methods to another context*, and *illustrating with examples*. The authors emphasize there may be interesting relationships between these and call for further investigation into these relationships. For this report, we present results from adapting the Mejía-Ramos et al. (2012) model to analyze mathematics faculty responses to a survey in which they described their own approaches to self-assessment in proof writing. As this is a preliminary report, the ultimate goal of the results from this survey is to design a tool faculty can provide to students to help prompt the use of proof comprehension in self-assessment and to improve overall proof writing skills. We present and describe this adapted model in the Results section and provide examples of prompts meant to provoke these types of self-assessment via proof comprehension strategies.

Methodology

Several studies about proof comprehension and self-assessment emphasize the importance of the collaboration of mathematicians and mathematics educators, as mathematicians contribute an expertise in proof reading, an understanding of what is reasonable for students to understand, and they are more likely to embrace research results when they are actively involved (Cooley et al., 2024; Mejía-Ramos et al., 2012; Yang & Lin, 2008). Our current project takes this perspective, as we surveyed tenure-track and tenured mathematics faculty at a Hispanic Serving Institution (HSI) who regularly teach or have recently taught a mathematics course involving proof writing. We note the researchers on this project include a Hispanic male undergraduate student and a white female faculty member. The Institutional Review Board (IRB)-approved online survey was conducted anonymously, and out of the 14 faculty members recruited via email in the department at the institution, seven volunteered to participate. The 22 item survey (20 of which were open-ended) included two mathematical statements for faculty to consider: (a) For $a, b \in \mathbb{R}$, prove that $|a+b| \leq |a|+|b|$, and (b) For all integers a, b, c, if $a, b,$ and c are odd, then the equation $ax^2+bx+c=0$ has no solution that is a rational number. Faculty were asked questions about their own proof writing approaches for these statements, proof self-assessment strategies, other reflection questions, and then were prompted to assess a sample student proof for each of the statements. The data collected from the survey was analyzed first using open-coding (Richards & Hemphill, 2018) techniques in several passes by the researchers, and these codes were negotiated throughout. Trends emerged regarding ways that mathematicians approach self-assessment in their proof writing and the criteria faculty use for themselves and their students to evaluate the success of a proof. Many of these codes and subcodes resembled the model presented in Mejía-Ramos et al. (2012), and so we re-analyzed our data through this lens. For this report, we focus on Faculty A’s responses to two open-ended questions from the survey that prompted participants to describe their self-assessment strategies and evaluation criteria in proof writing. We are using the results from this data as guidance to design a tool students may use to help with self-assessment of their proof writing. We plan to eventually interview students while using this tool to assess its effectiveness.

Results

Below, we provide the types of assessment presented in Mejía-Ramos et al. (2012), adapted for self-assessment as it was used in our analysis. We note that 1-3 below relate to *local* and 4-6 relate to *holistic* types of self-assessment (we omit the adaptation of the seventh type of assessment from Mejía-Ramos et al., (2012), *illustrating with examples*, as we found this code came up primarily in relation to constructing a proof rather than self-assessment).

1. *Use of terms and statements* - To assess clarity and use of key terms and statements, including explanations of these
2. *Logical status of statements and proof framework* - To assess validity of statements in the proof and the logical relationship between them and the given proposition
3. *Justification of claims* - To assess comprehensibility of the chaining logic in the proof and use of other assertions and assumptions
4. *Presentation via high-level ideas* - To assess presentation and communication of the main ideas of the proof and the overarching approach
5. *Logical flow with the modular structure* - To assess the logical flow and proof structure in terms of its main components/modules and the logical relationship between them
6. *Other Contexts and Criteria* - To assess "completeness" by comparing and adapting ideas from other proving tasks to evaluate the methods or procedures of current proof

In Table 1 below, we provide Faculty A's responses to two open-ended questions about both statements on the survey (listed in Methodology). Note as the survey was conducted anonymously, we do not have demographic information of Faculty A. We select this participant's responses for this preliminary report because they are thoughtful, representative of the trends found among the seven surveyed faculty, and they illustrate each of the self-assessment types in our adapted model. In this table, we utilize the numbering system above to connect each statement made by the faculty member to these self-assessment types. We describe how these results can be leveraged pedagogically in the Discussion and Conclusion section.

Table 1. Faculty A's responses to selected open-ended questions from survey.

Questions (open ended)	Responses from Faculty A with superscripts connected to adapted model
How did you assess your own proof as you were writing/sketching it? (What things do you look for in your own proof to decide if it is "going in the right direction"?)	Statement (a): [Can I justify the next line given the previous lines?][3] [Am I making unwarranted assumptions (i.e., if pressed could I explain my reasoning to a skeptic?)?][1,2] [Will I understand and agree with my own proof later?][4] Statement (b): [Are there specific details I can leverage or latch on to or are my givens still too vague.][1,3] [What can I do to get a handle on them?][3] [Can I simplify this to look like a previous or easier problem?][6]
Once you feel you are done writing a proof of this statement, how do you assess it for correctness? What things do you look for in your own proof to decide if it is valid?	Statement (a): [For a proof by cases like this one: Are my cases exhaustive?][5] [Am I making too many assumptions?][2] [Can I follow it from one line to the next using sound and valid logic?][3] [Would I be able to explain myself rigorously to another (vs. just making a persuasive argument)?][4] Statement (b): [Logical flow][3,5], [validity][1,2,4], [proof structure (in this case, exhaustive cases and good proof by contradiction structure).][5,6]

When specifically asked about the criteria that Faculty A would use to assess their own proof work, they said, "Validity, soundness, clarity, comprehensibility. Depending on use case, I might add explanatory power if it's meant to demonstrate reasons behind a mathematical principle, or brevity if searching for ‘elegant’ proofs.” This is consistent with Beumann and Wegner (2018) in that self-assessment should include a consideration for “personal taste,” depending on the writer and the intended reader. For example of analysis, when Faculty A said, “Are there specific details I can leverage or latch on to or are my givens still too vague[?],” we interpret this as the faculty member evaluating if they have leveraged all of the given conditions appropriately in their proof, specifically to improve clarity. Thus, this question is classified as self-assessment types 1 and 3 and is focusing on the faculty’s criterion of “clarity”.

Discussion and Conclusion

Recall our research questions regarding mathematicians’ proof self-assessment strategies and ways to incorporate proof comprehension in self-assessment. Using the Mejía-Ramos et al. (2012) model adapted for proof self-assessment, we analyzed seven faculty responses to open-ended survey questions about their self-assessment process and criteria. We used these responses and the model to guide the development of a draft of a self-assessment tool using proof comprehension strategies. Below, we provide examples of prompts that target these self-assessment types that students may use as a tool in proof validation and self-evaluation.

1. *Use of terms and statements* - Did I consider/address all conditions in the proposition? Did I use each concept from the proposition with the appropriate assumptions?
2. *Logical status of statements and proof framework* - What is the purpose of each statement? What justifies making each assumption? How does each statement contribute to the proof framework I am using?
3. *Justification of claims* - Can I justify this line given the previous lines in my proof? Am I leaving out any information that is necessary for a later implication?
4. *Presentation via high-level ideas* - Conceptually, what is the “big idea” of my proof and is it communicated clearly? Is the overarching approach of my proof sound?
5. *Logical flow with the modular structure* - What are the main modules or components of my proof? What is the purpose of each of these components? Do the components work together logically as intended for the type of proof being used?
6. *Other Contexts and Criteria* - Does my proof resemble a familiar proof of a similar statement (in structure)? Does this proof fit the relevant criteria for its intended use or evaluation? Could the intended reader understand and follow my proof?

We suggest the use of these prompts in undergraduate proof writing courses as an activity to encourage self-assessment and opportunities for self-reflection on proof writing. We believe these may improve student self-efficacy in proof writing and communication about evaluation criteria. Future research could investigate the use of this tool in classrooms, specifically in small group proving tasks, as Beumann and Wegner (2018) suggest. We are still refining this tool through deeper analysis of our faculty data, including the consideration of faculty criteria when evaluating student proofs. Our main goal is to assist students with the proof self-assessment process through the lens of proof comprehension. As part of our larger project, we plan to collect data from student interviews using the self-assessment tool to investigate its possible influence on student self-efficacy and proof writing skills. Obtaining other perspectives from additional undergraduate mathematics faculty would also help in improving this adapted model and tool.

Questions for the Audience

1. What self-assessment strategies or activities have you used in proof writing and what obstacles do you feel your students face (or would face) in this self-assessment?
2. What suggestions do you have to improve our adapted model and prompts?

References

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory into Practice*, *48*(1), 12-19.

Azrou, N. (2015, February). Proof writing at undergraduate level. In *CERME 9-Ninth Congress of the European Society for Research in Mathematics Education* (pp. 79-85).

Beumann, S., & Wegner, S. A. (2018). An outlook on self-assessment of homework assignments in higher mathematics education. *International Journal of STEM education*, *5*, 1-7.

Chamberlain Jr, D., & Vidakovic, D. (2021). Cognitive trajectory of proof by contradiction for transition-to-proof students. *The Journal of Mathematical Behavior*, *62*, 100849.

Chamberlain, D. J. Jr. (2017). *Investigating the development of proof comprehension: The case of proof by contradiction*. [Doctoral Dissertation, Georgia State University].

Cooley, L., Dorfmeister, J., Miller, V., Duncan, B., Littmann, F., Martin, W., ... & Yao, Y. (2024). The PRIUM qualitative framework for assessment of proof comprehension: a result of collaboration among mathematicians and mathematics educators. *ZDM–Mathematics Education*, 1-14.

Conradie, J., & Frith, J. (2000). Comprehension tests in mathematics. *Educational Studies in Mathematics*, *42*, 225-235.

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

Hodds, M., Alcock, L., & Inglis, M. (2014). Self-explanation training improves proof comprehension. *Journal for Research in Mathematics Education*, *45*(1), 62-101.

Kemp, A., Chamberlain Jr., D., Cooley, L., Miller, V., Vidakovic, D. (2020). Student Self- and Simulated Peer-Evaluation of Proof Comprehension: Tina. In: Karunakaran, S. S., Reed, Z., & Higgins, A. (Eds.), *Proceedings of the 23rd Annual Conference on Research in Undergraduate Mathematics Education* (pp. 303-310).

Ko, Y., & Knuth, E. (2013). Validating proofs and counterexamples across content domains: Practices of importance for mathematics majors. *The Journal of Mathematical Behavior, 32*(1), 20–35.

Lew, K., & Mejía-Ramos, J. P. (2019). Linguistic conventions of mathematical proof writing at the undergraduate level: Mathematicians' and students' perspectives. *Journal for Research in Mathematics Education*, *50*(2), 121-155.

Mejía-Ramos, J. P., Fuller, E., Weber, K., Rhoads, K., & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics*, *79*, 3-18.

Mejía-Ramos, J. P., & Weber, K. (2019). Mathematics majors' diagram usage when writing proofs in calculus. *Journal for Research in Mathematics Education, 50*(5), 478–488.

Melhuish, K., Fukawa-Connelly, T., Dawkins, P. C., Woods, C., & Weber, K. (2022). Collegiate mathematics teaching in proof-based courses: What we now know and what we have yet to learn. *The Journal of Mathematical Behavior*, *67*, 100986.

Poehner, M. E. (2012). The zone of proximal development and the genesis of self-assessment. *The Modern Language Journal*, *96*(4), 610-622.

Powers, R. A., Craviotto, C., & Grassl, R. M. (2010). Impact of proof validation on proof writing in abstract algebra. *International Journal of Mathematical Education in Science and Technology*, *41*(4), 501-514.

Richards, K. A. R., & Hemphill, M. A. (2018). A practical guide to collaborative qualitative data analysis. *Journal of Teaching in Physical Education*, *37*(2), 225-231.

Selden, A., & Selden, J. (2015). Validations of proofs as a type of reading and sense-making. In: Beswick, K., Muir, T., and Fielding-Wells, J. (Eds), *Proceedings of the 39th Conference of the International Group of Psychology of Mathematics Education* (Vol. 4, pp. 145-152). Hobart, Australia: Psychology of Mathematics Education (PME).

Selden, A., & Selden, J. (2017). A comparison of proof comprehension, proof construction, proof validation and proof evaluation. In: Robin Göller, Rolf Biehler, Reinhard Hochmuth, Hans-Georg Rück (Eds.), *Proceedings of the Conference on Didactics of Mathematics in Higher Education as a Scientific Discipline* (pp. 339-345).

Stallings, V., & Tascione, C. (1996). Student self-assessment and self-evaluation. *The Mathematics Teacher*, *89*(7), 548-554.

Stylianides, G. J., Stylianides, A. J., & Weber, K. (2017). Research on the teaching and learning of proof: Taking stock and moving forward. *Compendium for Research in Mathematics Education*, 237-266.

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.

Weber, K., & Tanswell, F. S. (2022). Instructions and recipes in mathematical proofs. *Educational Studies in Mathematics*, *111*(1), 73-87.

Yang, K. L., & Lin, F. L. (2008). A model of reading comprehension of geometry proof. *Educational Studies in Mathematics*, *67*, 59-76.