**Improving the Quality of Proofs for Pedagogical Purposes: A Quantitative Study**
Contributed Research Report

In university mathematics courses a primary means of conveying mathematical information is by mathematical proof. A common suggestion to realize learning goals related to proof is to increase the quality of the proofs that we present to students. The goal of this paper is to investigate evidence related to the question: What changes to a proof do mathematicians believe will improve the quality of a proof for pedagogical purposes?

We present quantitative findings that corroborate hypotheses generated by a qualitative study of features of proofs that mathematicians find pedagogically valuable. Our work examines hypotheses related to typesetting, brevity, and the framework of a proof. One of our findings suggests that there is not a consensus among mathematicians what level of justification is desirable or necessary for the purposes of teaching undergraduates, though there may be common themes in the warrants they give for the level chosen.
*Key words*: proof evaluation, mathematicians, proof revision, quantitative study.

## 1. Introduction

In university mathematics courses, a primary means of conveying mathematical information is by mathematical proof. We hope that these proofs can convince students that a theorem is true, illustrate why a theorem is true, or illustrate new methods of reasoning (e.g., de Villiers, 1990; Hanna, 1990; Hanna & Barbeau, 2008). Unfortunately, both empirical and anecdotal evidence indicate that these learning goals often are not realized. Selden and Selden (2003) argue that as undergraduates do not have the ability to differentiate valid and invalid proofs (Selden & Selden, 2003; Weber, 2010), they cannot be gaining legitimate mathematical conviction from the proofs that they read. Many researchers report that students find the proofs they read to be confusing or pointless (e.g., Harel, 1998; Hersh, 1993; Porteous, 1986; Rowland, 2001).

A common suggestion to improve this situation is to increase the quality of the proofs that we present to students. For instance, some researchers argue that proofs should be more closely tied to informal arguments (Hersh, 1993), make explicit the proof's overarching structure while suppressing logical details (Leron, 1983), or illustrated with a carefully chosen generic example (Rowland, 2001). We attempt to build on this work by addressing the question: What types of modifications do mathematicians believe will improve the quality of a proof for pedagogical purposes?

Last year, at the 2010 RUME Conference, we presented the results of a qualitative study in which we asked eight mathematicians to revise two proofs intended for a calculus course for second- or third-year mathematics majors. We suggested a number of features in proofs that mathematicians find pedagogically valuable (see Lai & Weber, 2010). However, due to small sample sizes and the qualitative nature of our study, we qualified our findings as "grounded hypotheses". The goal of this study is to test these hypotheses with a larger number of mathematicians. Specifically, we aim to test the following hypotheses:

**(H1)** A proof for undergraduates can be improved if a hypothesis and conclusion statements are added to the proof that make explicit the proof framework (in the sense of

Selden & Selden, 1995) being employed. (In many undergraduate proofs, these are not explicitly stated and the proof framework is implicit).

**(H2)** Emphasizing important equations in a proof via typesetting will improve the quality of the proof because it will make clear the proof's main ideas.

**(H3)** Adding extra justification to support an assertion can improve the clarity of a proof if that justification might be difficult for a student to infer on their own.

**(H4)** Including unnecessary irrelevant computations or assumptions in a proof will make the proof worse since this will unnecessarily lengthen the proof and confuse students.

## 2. Theoretical assumptions

This work is based on three theoretical assumptions. First, understanding teachers' pedagogical beliefs is essential for modifying teachers' behavior (e.g., Aguirre & Speer, 1996). It follows that understanding what mathematicians believe constitutes a good proof for pedagogical purposes is necessary if we want to change the way that proofs are presented in university classrooms. Second, as experienced practitioners, mathematicians are a useful source of pedagogical content knowledge (Alcock, 2010). Consequently, mathematicians' views on how proofs might be improved are useful considerations for mathematics educators to consider. Third, small-scale qualitative studies are essential for developing grounded hypotheses in mathematics education research; however, these hypotheses need to be rigorously tested.

## 3. Methods

To test the hypotheses, we conducted an internet-based study. To seek participants, we sent e-mail requests to 28 mathematics departments inviting their faculty members, post-docs, and Ph.D students to participate in our study, providing a link to the website that they could click on to participate in the study. 110 mathematicians chose to participate. Methodological details about measures we took to insure the validity of this study, as well as empirical evidence that this approach is valid, are similar to the methods and arguments given in Inglis and Mejia-Ramos (2009).

In this study, participants were then shown the "master proof" below and told they would be asked whether changes to the proof would make it "less or more understandable to second or third year undergraduate students".

**Proposition.** *If $f : \mathbb{R} \to \mathbb{R}$ is differentiable and $f'(x) > 0$ for all $x \in \mathbb{R}$, then $f$ is injective.*

*Proof.* Let $x_1, x_2 \in \mathbb{R}$, where $x_2 > x_1$. The Mean Value Theorem implies there exists $x_3 \in [x_1, x_2]$ such that $f'(x_3) = \frac{f(x_2)-f(x_1)}{x_2-x_1}$. Since, by hypothesis, $f'(x_3) > 0$ and $x_2 - x_1 > 0$, then $f(x_2) - f(x_1) = f'(x_3)(x_2 - x_1) > 0$. Therefore $f(x_2) \neq f(x_1)$. □

Figure: "master proof" shown to participants.

The participants were then shown a screen with the master proof on top and a modified version of the master proof with the modifications in blue at the bottom and were asked to judge whether the changes made the proof "significantly better", "somewhat better", "the proofs were the same", "somewhat worse", or "significantly worse" (which we coded as 2, 1, 0, -1, or -2 respectively). This process was repeated five times with five different modified proofs. The order in which each proof was received was randomized.

**(M1)** We presented a proof where we added the sentence, "To show f is injective, we must show that $f(x_1) \neq f(x_2)$ after the first sentence of the master proof and the sentence "It follows that f is injective" as the last sentence of the proof. If H1 is correct, the participants should judge M1 to be an improvement over the master proof.

**(M2)** The formulas, $f'(x_3) = f(x_2) - f(x_1)/(x_2 - x_1)$ and $f(x_2) - f(x_1) = f'(x_3)(x_2-x_1) > 0$, were re-formatted to appear centered as their own lines. If H2 is correct, the participants should judge M2 to be an improvement over the master proof.
**(M3)** The last sentence of the proof was re-written as "Since $f(x_2) - f(x_1) \neq 0$, $f(x_2) \neq f(x_1)$". This added an extra justification that had previously been implicit to the proof. If H3 is correct, participants should judge M3 to be an improvement over the master proof.
**(M4)** We added the phrase "so $x_2 - x_1 = f(x_2) - f(x_1)/f'(x_3)$" immediately before the sentence beginning with "Since". While this inference is correct, it is not useful in the proof. If H4 is correct, participants should judge M4 as worse than the master proof.
**(M5)** We added the phrase "f is a real valued function" after the phrase "Since, by hypothesis". This assumption was not relevant to the subsequent argumentation. If H4 is correct, participants should judge M5 as worse than the master proof.
Finally, we gave the participants the option of commenting on why they made the judgment that they did.

## 3. Results

A repeated measures ANOVA revealed a main effect based on the modifications the participants' received ($F(327, 4) = 231.7$, $p<0.001$), meaning participants did not judge all modifications to be of equal quality. A summary of the results is given in the table below.

| Condition | Mean score | # participants who thought proof was better | # participants who thought proof was worse |
|---|---|---|---|
| M1 | $1.29^*$ | 97 | 4 |
| M2 | $1.05^*$ | 88 | 2 |
| M3 | 0.02 | 41 | 40 |
| M4 | $-1.66^*$ | 6 | 98 |
| M5 | $-1.12^*$ | 7 | 94 |

$^*$ Indicates a mean score statistically different than zero with $p<0.001$.

These results confirm the predictions based on H1, H2, and H4. However, participants' responses for M3 fail to confirm H3 (the hypothesis that an extra justification will help students). Among the participants, 41 mathematicians thought adding the extra justification in M3 to the proof would make it better, with some giving reasons such as "I can imagine students being confused by the last step and this change would make it clearer". The 41 who thought the change lowered the quality of the proof gave responses that they thought students should make this inference easily, that students should be pushed to make these inferences on their own, or that the inference added was worded poorly. This suggests there is not a consensus among mathematicians for what level of justification is desirable or necessary for the purposes of teaching undergraduates.

## 4. Significance and future research

These results confirm several of the hypotheses proposed by the authors based on a qualitative study at last year's RUME conference (Lai & Weber, 2010). In particular, mathematicians believe that, for purposes of pedagogy, brevity is a desirable attribute of a proof but adding proof frameworks improves their quality. Formatting a proof by centering important equations also improves their quality. An interesting future research question is if the changes the mathematicians endorse would improve students' comprehension of proofs.

**References**


Alcock, L. (2010). Mathematicians perspectives on the teaching and learning of proof. *Research in Collegiate Mathematics Education*, 7. 73-100.

Aguirre, J. and Speer, N. (1996). Examining the relationship between beliefs and goals in teacher practice. Paper presented at the annual meeting of the American Educational Research Association, New York, April 8-12, 1996.

de Villiers, M. D. (1990). The role and function of proof in mathematics. *Pythagoras, 24*, 17-24.

Hanna, G. (1990). Some Pedagogical Aspects of Proof. *Interchange*, 21(1), 6-13.

Hanna, G. & Barbeau, E. (2008). Proofs as bearers of mathematical knowledge. *ZDM, 40*, 345-353.

Harel, G. & Sowder, L. (1998). Students proof schemes. *Research in collegiate mathematics education,* 3, 234-282.

Hersh, R. (1993). Proving is convincing and explaining. *Educational Studies in Mathematics*, *24*, 389-399.

Inglis, M. & Mejia-Ramos, J.P. (2009). The effect of authority on the persuasiveness of mathematical arguments. *Cognition and Instruction*, 27, 25-50.

Leron, U. (1983). Structuring mathematical proofs. *American Mathematical Monthly*, 90(3), 174-184.

Porteous, K. (1986). Children's appreciation of the significance of proof. In *Proceedings of the Tenth International Conference for the Psychology of Mathematics Education*, 392-397. London, England.

Rowland, T. (2001). 'Generic proofs in number theory.' In S. Campbell and R. Zazkis (Eds.) *Learning and teaching number theory: Research in cognition and instruction*. (pp. 157-184). Westport, CT: Ablex Publishing.

Selden, A. & Selden, J. (2003). Validations of proofs written as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education, 36*(1), 4-36.

Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics, 29,* 123-151.

Weber, K. (2010). Mathematics majors' perceptions of conviction, validity, and proof. *Mathematical Thinking and Learning*, 12, 306-336.