

The State of Research on Proof and Proving at RUME Conferences from 2018-2023:
A Systematic Literature Review

Amanda Lake Heath
Middle Tennessee State University

Rosina Andrews
Middle Tennessee State University

Jordan E. Kirby
Francis Marion University

Sarah K. Bleiler-Baxter
Middle Tennessee State University

Proof and proving are critical to the work of mathematicians, and in turn research on the teaching and learning of proof has historically been one of the larger strands of research presented at annual RUME conferences. In this project we systematically reviewed the RUME conference papers on mathematical proof from 2018-2023 to provide a picture of (a) the most prevalent topics of proof research in RUME, and (b) the data sources drawn upon in proof research methodologies in RUME. This review revealed the most frequent primary topics of RUME proof research to be proof instruction, assessment and feedback, and proof conceptions. We also noticed an overwhelming proportion of RUME proof research papers rely on qualitative data sources. Our continued work on this project will allow us to examine the impact research RUME has on the larger body of knowledge of undergraduate proof research.

Keywords: proof, research, literature review, methodology

Within the past six years, two major reviews of research on the teaching and learning of proof have been published (i.e., Stylianides et al., 2017; Stylianides et al., 2023), both casting a wide net and synthesizing research across K-16 mathematical contexts. In the first review, a contributed chapter in the Compendium for Research in Mathematics Education, Stylianides et al. (2017) examined the field of research on the teaching and learning of proof across K-16 mathematics education and classified the proof literature into three research perspectives: proof as problem solving, proof as convincing, and proof as a socially-embedded activity. In the second review, in a recent issue of ZDM-Mathematics Education, Stylianides et al. (2023) again examined the field of research on teaching and learning of proof across K-16 mathematics education, but this time classified the research according to the various didactic relationships emphasized between student, teacher, and mathematical content. We aim to add to the literature review work of these authors by (a) focusing exclusively on recent research accepted and published in the RUME proceedings, (b) taking a more emergent approach to identifying topics studied by proof researchers, and (c) attending to the methods utilized by proof researchers.

The reasons for isolating our literature search to RUME conference proceedings are threefold. First, by focusing on the RUME conference proceedings, we can better understand how this community of scholars is approaching proof research and empower the community to recognize its strengths and areas for growth. Second, we argue the RUME conference proceedings can indicate the most up-to-date trends in research, as the conference accepts contributed, preliminary, and theoretical proposals, providing a space for researchers to share early ideas and research-in-progress. Third, confining our literature search to RUME conference papers will enable us to later observe patterns in how proof research presented at the RUME conference evolves (or fails to evolve) into eventual journal publication, providing greater insight into the impact of the community's work.

Our goals in this review are to, in RUME, (1) identify trends in topics of proof research and (2) identify trends in methods of proof research. Specifically, our research questions are:

1. When RUME researchers study proof, what topics are they investigating? (In other words, what are they studying as it relates to proof?)
2. When RUME researchers study proof, what data sources are they drawing upon? (In other words, through what mechanism(s) are they telling the story of their research?)

Theoretical Framing: Advocating for Variability of Topics and Methods

In this research, we come from a pragmatic perspective, which assumes “knowledge guides method while method also guides knowledge” (Paul, 2005, p. 46). In this way, we seek to understand both the topics proof researchers are attending to in their research and the methods they are using to do so. We believe that variability in both the topics of investigation and methods of investigation can lead to a field of research that is more robust and has greater impact on practice. We draw from Stoecker and Avila (2021), who describe the benefits of addressing complexity of topics through variability of methods:

From [a mixed-methods research] perspective, quantitative and qualitative methods are complementary, each compensating for the other’s weaknesses in studying complexity. [...] Greene (2007) and Clark and Ivankova (2016) believe that different methods can produce different kinds of information, such that one method can initiate research questions for another method, or expand a research project and lead to broader conceptual development [...] and] even advocate that mixed methods research can promote social justice. (Stoecker & Avila, 2021, pp. 627-628)

In particular, we recognize different data sources have a different ability to tell a story. If we, as RUME researchers, have a preference toward certain data sources we use to make conclusions, we may be getting a limited perspective on a given proof topic. For example, consider video data from a classroom setting versus in a task-based interview setting. The classroom video data can help us better understand the proof learning contexts we are most interested in influencing, while the interview data provides a potentially cleaner and more focused picture of proof topics. Similarly, survey data has a potentially wide (but not deep) reach while interview data offers a potentially deep (but not wide) reach. Variety of methods is critical to the impact, robustness, and balanced perspective of a field of study (Stoecker & Avila, 2021).

Moreover, exploring the topics of study within a field (i.e., proof research), and in this case within a professional community (i.e., RUME), will allow us to identify areas of proof research that are over- or under-represented. Given the recent reviews of research on proof and proving we summarized above, we believe it will benefit the RUME community to compare the results of this literature review to those of other reviews. Previous proof reviews (Stylianides et al., 2017; Stylianides et al., 2023) were focused on proving literature related to teaching and learning in K-16 broadly, or on the teaching proof in undergraduate settings (Melhuish et al., 2022) and thus future research could benefit from comparing the RUME-specific topics of study from this review to the findings from previous broader reviews.

Methods

In this study examined proof research published in the RUME conference proceedings during 2018-2023. After collecting all contributed, preliminary, and theoretical reports from these years, we identified reports related to proof by a word search for “proof” or “proving” in both the paper titles and abstracts. The four authors sorted these reports into three categories: Proof Inquiries, Proof Settings, and Unrelated. Proof Inquiries include research about the teaching, learning, or

professional practices of proof or research regarding proof-related concepts/topics conducted in a proof-based course context. Proof Settings refers to research that occurs in a proof course but was not investigating topics specific to proof. Finally, reports classified as Unrelated are those reports with the words “proof” or “proving” in the text of the title or abstract, yet these words were not used in the sense of mathematical proof. All reports were sorted by at least two authors who met to reach consensus on any disagreements. We excluded the 2021 RUME Reports due to the difference in that year’s call for research to be conducted primarily by early career scholars. Moreover, our analysis of the 2023 proceedings is ongoing. We describe the methodological approach used to answer each research question below.

RQ1 Methods: Topics in RUME Proof Research

To identify the topics of RUME Proof research, we took an emergent approach to analysis. First, all four authors independently described the study focus of each report in their own words. We collapsed our individual descriptions into a list of topic codes through a card sort activity. We then each independently coded reports using these topic codes by assigning weights to the most relevant topics (3: Most Relevant Topic; 2: Very Relevant Topic; 1: Relevant Topic). Although each researcher was required to assign a most relevant (3) code to each study, topics weighted with a (2) or (1) were optional. Each proceedings year was coded by two authors, and although we did not seek to reach exact agreement, we discussed our codes in pairs and sought to understand each other’s rankings and potentially altered our codes following further discussion.

We analyzed our topic codes by first summing the weighted topic codes of both coders on each article, so the topic code for each report ranged from 0 to 6. Our analysis proceeded in two ways: (1) we considered the topic codes with maximum sum value as the “primary topic” for each report, and (2) we observed trends in overall topic prevalence, regardless of weight, by considering the frequency of topics by code sum greater than or equal to one.

RQ2 Methods: Data Sources in RUME Proof Research

We first developed a list of data source codes by consulting textbooks and publications regarding educational research methods (e.g., Creswell, 2013; Ottinger, 2019; Merriam & Tisdell, 2016). Although our research team originally set out to classify research reports according to a set list of categories, after many efforts to precisely define our data sources, we determined this goal was much more complicated than expected. Upon further reading, we learned we are not the first mathematics education researchers who have encountered such difficulties. Schoenfeld (2010) described his efforts to create a taxonomy of research methods,

My original intention [...] was to provide a selective overview of some relevant categories of research methods, and to raise some issues about their use. This is by no means a straightforward task. (p. 495).

After reaching a conclusion that, like Schoenfeld (2010), we could not create a list of well-defined, mutually exclusive sources of data, we decided to proceed with our list of codes, treating each code as independent despite potential overlap. The categories for analysis were theoretical reports, literature reviews, qualitative data sources in the form of interviews, observations, artifacts, and surveys, and quantitative data sources in the form of surveys and assessments. The four researchers met in teams of two to discuss and resolve any disagreements.

Preliminary Findings

Across the RUME conference proceedings in the last five years (2018-2023), there have been 736 preliminary, contributed, and theoretical reports. Our search for reports regarding proof

yielded 88 total reports: 18 reports from 2018, 15 reports from 2019, 20 reports from 2020, 18 reports from 2022, and 18 reports from 2023. Broadly, proof papers have composed about 12.3% of preliminary, contributed, and theoretical reports published in the RUME proceedings in the last five years (excluding 2021). We report preliminary findings from the 70 proof-focused reports from 2018-2022, as analysis of the 2023 proceedings is ongoing.

RQ1 Findings: Topics in RUME Proof Research

Across the RUME proceedings for which our research team has completed analysis, the most frequent primary topic in the RUME proof reports was proof instruction and feedback (12/70, 17.1%), followed closely by proof conceptions (views/beliefs about proof; 11/70, 15.7%). In contrast, the topic least prevalent as a primary focus was the transition to proof (2/70, 2.9%). Analysis of the primary topics of each report also revealed the, unsurprising, emergence of investigations of proof and online learning and technology in 2022, likely due to the transition to online instruction of many university courses during the COVID-19 pandemic. Figure 1 illustrates the prevalence of the primary topics of RUME proof research across 2018-2022.

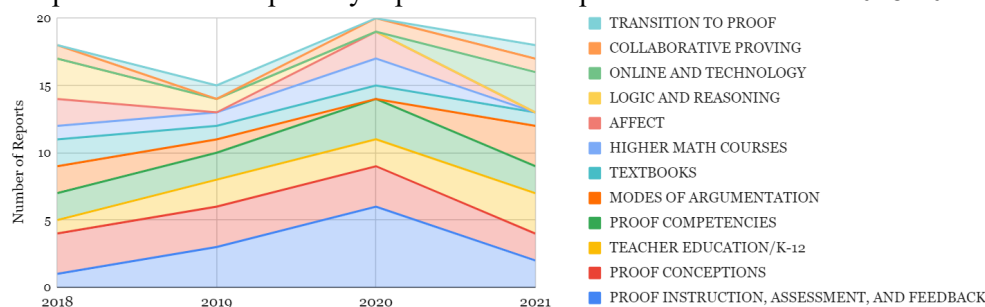


Figure 1: Primary Topics of RUME Conference Proof Research 2018-2022

By viewing the data, instead, with respect to the topics labeled any level of relevancy for each report rather than only the primary topic, we also observe trends in potential secondary topics and/or contexts of the RUME proof research. For example, although the transition to proof was least frequently the primary topic of investigation for the proof reports 2018-2022, it was the topic most frequently marked as relevant to the study in some capacity (i.e., at least one of the coders marked Transition to Proof as a 3, 2, or 1). Meanwhile, the investigation of textbooks appeared less frequently as a topic across all codes, yet when it was labeled as relevant this was often the primary topic for a report, so we can deduce textbooks are rarely a secondary or tertiary topic focus for proof research in RUME.

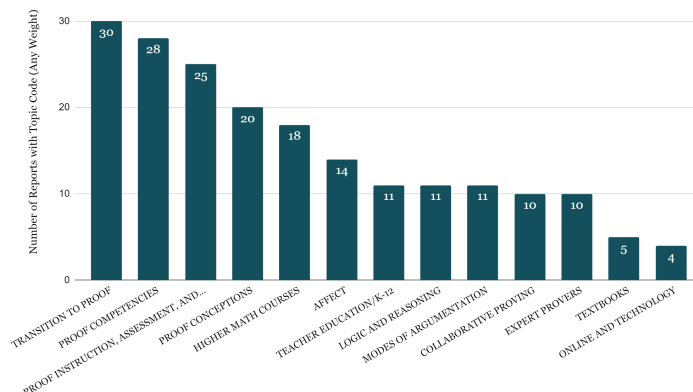


Figure 2: All Topics Marked Relevant (Any Weight) in RUME Proof Research 2018-2022

RQ2 Findings: Data Sources in RUME Proof Research

Because we encountered difficulties developing precise code definitions for data sources, we took suggestions from Schoenfeld (2010) in conjunction with methodological precedence from DiMartino and colleagues (2022) and classified RUME reports on proof as having drawn upon qualitative data, quantitative data, or both. This simple classification of reports revealed a drastic disparity between qualitative and quantitative data in RUME proof research. Two relevant theoretical reports did not employ any empirical data, so leaving 68 reports to consider data sources. Of these reports, 67 (98.5%) used some source of qualitative data, whereas only 7 (10.3%) used some form of quantitative data. Six reports used both qualitative and quantitative data, and therefore there was only one report which drew upon quantitative data alone. A final notable finding is the prevalence of interview research. Thirty-four (50%) of the proof reports using empirical data included interview data.

Discussion and Conclusion

After analyzing the 736 preliminary, contributed, and theoretical reports published in RUME conference proceedings in 2018-2023, we identified some of the most common topics of and data sources used in proof research within the RUME community. These preliminary findings indicate there is a strong presence of research regarding proof instruction, assessment, and feedback as well as proof conceptions, but there is much less research regarding collaborative proving, online learning and technology, and logic and reasoning. Much of the proof research ($n=30$, 42.8%) concerned the transition to proof, yet only two of these reports (2.8%) had the transition to proof as its primary topic. This finding indicates that there are many researchers conducting research within transition-to-proof courses but are not as frequently foregrounding students' transition to proof-based mathematics as their research focus.

With respect to methodologies and data used in RUME proof research, it is clear an overwhelming majority of reports used some form of qualitative data, and only one report used purely quantitative data. We speculate this absence of quantitative data may be due to the lack of reliable measurement tools relevant to the research of mathematical proof. Although we have not yet explored the reports in each topic in depth, we are interested to discover if RUME research regarding proof instruction, assessment, and feedback might be working toward developing measures which would allow for the collection of quantitative data.

In conclusion, we see several areas in which the RUME proof community could grow to assure there is a diverse presentation of topics and methodologies used to investigate proof at our conferences. In particular, there is room for more research on the transition to proof as a moment of mathematical development, and research on collaborative proving and proof learning and instruction through online mediums and technology. We also recognize the need for more quantitative data in order to complement the depth of qualitative data with widespread, generalizable claims. In moving this work forward, we intend to complete analysis on the 2023 proceedings, and then shift our focus to research journal publications on proof at the undergraduate level. This shift to journal publications will allow us to compare the distribution of research topics at RUME to the larger body of work in undergraduate proof research as well as track how research presented at RUME conferences evolves (or fails to evolve) into eventual journal publication. We pose the following questions for discussion with our audience:

1. What strategies for classifying proof research by methodology could best inform the field?
2. What research journals does our audience either seek publication in for proof research or look to for research on proving at the undergraduate level?

References

- Clark, V. L. P., & Ivankova, N. V. (2016). *Mixed methods research: A guide to the field*. SAGE Publications, Inc.
- Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. 4th Edition, SAGE Publications, Inc., London.
- Greene, J. C. (2007). *Mixed methods in social inquiry*. Jossey-Bass.
- Karunakaran, S. S., & Higgins, A. (Eds.). (2022). *Proceedings of the 24th annual conference on research in undergraduate Mathematics Education*. Boston, MA.
- Karunakaran, S. S., Reed, Z., & Higgins, A. (Eds.). (2020). *Proceedings of the 23rd annual conference on research in undergraduate mathematics education*. Boston, MA.
- 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.
- Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative Research: A Guide to Design and Implementation (4th ed.)*. San Francisco, CA: Jossey Bass.
- Ottinger, S. (2019). *Mathematical conjecturing and proving: the structure and effects of process characteristics from an individual and social-discursive perspective* [Doctoral Dissertation]. LMU München.
- Paul, J. L. (2005). *Introduction to the philosophies of research and criticism in education and the social sciences*. Upper Saddle River, NJ: Prentice Hall.
- Schoenfeld, A. H. (2010). Research methods in (mathematics) education. In L. D. English & D. Kirshner (Eds.), *Handbook of international research in mathematics education* (pp. 481-533). Routledge.
- Stoecker, R., & Avila, E. (2021). From mixed methods to strategic research design. *International Journal of Social Research Methodology*, 24(6), 627-640.
- Stylianides, G. J., Stylianides, A. J., & Moutsios-Rentzos, A. (2023). Proof and proving in school and university mathematics education research: a systematic review. *ZDM–Mathematics Education*, 1-13.
- Stylianides, G. J., Stylianides, A. J., & Weber, K. (2017). Research on the teaching and learning of proof: Taking stock and moving forward. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 237–266). Reston, VA: National Council of Teachers of Mathematics.
- Weinberg, A., Moore-Russo, D., Soto, H., & Wawro, M. (Eds.). (2019). *Proceedings of the 22nd annual conference on research in undergraduate mathematics education*. Oklahoma City, Oklahoma.
- Weinberg, A., Rasmussen, C., Rabin, J., Wawro, M., & Brown, S. (Eds.). (2018). *Proceedings of the 21st annual conference on research in undergraduate mathematics education*. San Diego, California.