TEACHERS' REACTIONS TO ROUTINE AND ALTERNATIVE PRACTICES FOR PRESENTING PROOFS: A SURVEY EXPERIMENT

Bukola Ake	Justin Dimmel	Patricio Herbst
University of Maine,Orono	University of Maine,Orono	University of Michigan
bukola.ake@maine.edu	justin.dimmel@maine.edu	pgherbst@umich.edu

We conducted a multimedia survey experiment with a nationally representative sample (n = 405) of secondary geometry teachers. Participants were shown storyboard depictions of instructional episodes and asked to rate the appropriateness of the (hypothetical) teacher's actions using a Likert-like response format. We analyzed participants responses using ANOVA. The purpose of the experiment was to investigate how secondary geometry teachers expect students to communicate when presenting proofs during class. Our results (1) replicated findings from a prior investigation of what teachers expect when students present proofs and (2) investigated how geometry teachers reacted to instructional practices that attempted to steer student presentations of proofs toward disciplinary communication practices.

Keywords: Mathematical Communication, Proof Transcriptions, Proof Presentation

Introduction

According to the National Council of Teachers of Mathematics: "Communication is a fundamental aspect of mathematics and mathematics education" (NCTM, 2000, p.60). The communication standard of NCTM emphasizes how crucial it is for students to express their mathematical reasoning to their peers and teachers using mathematical language. Mathematical communication enhances students' cognitive functioning by helping them articulate their ideas and strategy (Kosko & Wilkins, 2010; Lee, 2006).

One context where communication is central in mathematics classrooms is when students share their proofs with an entire class, an activity we refer to as *presenting proofs* (Dimmel & Herbst, 2020). Proof plays an important role in creating, developing, and communicating mathematical knowledge and is considered an important aspect of students' mathematical learning experiences (Ball & Bass, 2003; Herbst & Brach, 2006; Stylianides, 2019). How do secondary geometry teachers expect students to communicate when presenting proofs? This study investigated this question through a multimedia survey experiment. It is both a replication and an iteration of a prior study (Dimmel & Herbst, 2020) that investigated what teachers expect from students when they present proofs in geometry.

Background & Research Questions

Various modes of representation can be used when presenting proofs, including speaking, writing, diagramming, and gesturing (Stylianides, 2019). Several studies (e.g., Lai & Weber, 2014; Weber, 2004; Artemeva & Fox, 2011) have demonstrated how mathematicians employ various combinations of these modalities when presenting proofs in front of the class. For example, Weber (2004) examined the case of a professor presenting proofs in an introductory real analysis course. He found that the professor verbalized the proof's steps as they were written, offered commentary on the proof's development, and gave an overview of the proof before generating it.

Compared to the literature on students' written proofs, relatively little is known about how students present proofs (Kokushkin et al., 2022; Stylianides, 2019). Stylianides (2019) compared the oral arguments that secondary students made in front of the class with the written arguments they created during small group work for the same assertions. While this study explored how students communicate their proofs when required by an expert to do so, it neither provided any insight into the broader communication practices typical in mathematics classrooms, nor did it investigate teachers' expectations during such presentations.

To describe these broader communication practices, Dimmel & Herbst (2020) conducted a two-part, mixed methods study. First, they analyzed video records of geometry classrooms when students were asked to present proofs. They found that, in contrast to the multimodal practices of mathematicians, students engaged in mark-for-mark reproductions of previously completed proofs—an action described by the authors as *proof transcription*. A key aspect of this practice is that proofs were not expected to be evaluated as mathematical arguments until the transcription was completed. Second, they conducted a multimedia survey experiment in which hypothetical episodes of geometry instruction, represented by cartoon storyboards, were shown to a sample (n = 60) of midwestern secondary mathematics teachers. The storyboards showed students transcribing proofs and depicted different actions a teacher might take in response. Their study showed that the preferred action was no action – i.e., secondary mathematics teachers recognized proof transcription as an acceptable means for students to present proofs (Dimmel & Herbst, 2020). That teachers recognized proof transcription as routine is significant, because it suggests that there are limited opportunities for students to develop fluency with the multimodal communication practices for presenting proofs.

Our study builds on this prior work in two ways. One, we replicated the statistical tests described by Dimmel & Herbst (2018, 2020) with a larger, nationally representative sample (n = 405) of secondary geometry teachers. The replication study investigated whether this representative sample of geometry teachers recognized proof transcription as a routine instructional practice during proof presentation. Two, this study used two additional storyboards created to depict instructional actions that attempted to steer students toward developing multimodal communication skills during their proof presentations in geometry classrooms. While the findings from the study conducted by Dimmel & Herbst (2020) offer some insight into secondary mathematics teachers' expectations regarding how their students' present proofs, the design of the storyboards leaves open the possibility that teachers were merely responding or attending to teaching actions that represented a departure from other classroom instructional norms—e.g., an expectation of how students should be treated when they are in socially vulnerable situations such as when sharing their written work at the board, rather than attending to the communication practice depicted in the instructional scenario. Therefore, this present study provided a different way of scaffolding expert communication practices by designing storyboards targeting alternative communication practices designed to steer students away from simply transcribing their written proofs toward a multimodal presentation of their proofs. These two storyboards were designed to investigate how teachers would react to instructional episodes that require that students' modes of communication approximate the disciplinary practices of mathematical communication rather than an unreflective transcription of proofs (e.g., giving a verbal overview of the plan for the proof before writing it on the board). We thus operationalized our overarching question about how students learn discipline-specific mathematical communication practices into two research questions:

How do secondary geometry teachers react to instructional scenarios that either foster or condone the hypothesized norm of proof transcription when students present proofs at the board?

How do secondary geometry teachers react to instructional scenarios that create opportunities for students to engage in approximations to disciplinary communication practices?

Theoretical Framework: Instructional Situations and Normative Actions

We conceptualize teaching as a set of socially situated practices that occur in *instructional situations* (Herbst, 2006)—i.e., "identifiable segments of instruction that are organized around specific kinds of mathematical work, such as doing proofs in geometry or solving equations in algebra" (Dimmel & Herbst, 2020, p.7). Instructional situations are identifiable by their normative actions. Norms are familiar ways of observing, believing, evaluating, and behaving in an environment (Goodnough, 1971). They are defined in terms of the features of a social situation that not only regularly occur but that participants expect to occur (Garfinkel, 1963; Herbst & Chazan, 2003). The sociological view that people have unspoken expectations of how things should unfold has its roots in everyday experiences. For example, there are unspoken norms regarding how people stand in line when boarding a plane even though there are no explicitly written rules about the order in which people are to stand.

Instructional situations provide a way of framing the routine instructional activities of teachers and students in those situations. Thus, normative instructional activities, though invisible, can be investigated by observing how teachers react to scenarios that represent departures from those norms (Herbst & Chazan, 2015). The present study used the breaching experiment approach (Garfinkel, 1963) to investigate the normative ways secondary geometry teachers expect students to present proofs at the board. Specifically, we analyzed teachers' reactions to scenarios that represented departures from the norm of proof transcription.

Methods: Multimedia Survey Experiment with Planned Comparisons

We used a multimedia survey experiment with planned comparison between participants randomly assigned to different storyboard conditions. These different conditions were designed to probe teachers' responses to instructional actions that either interfered (treatment) or did not interfere (control) with student presentations of proofs. We used storyboard representations, rather than video segments of real-world classrooms, so that we could create hypothetical classroom vignettes that realized the specific instructional actions we wanted to investigate (Dimmel & Herbst, 2020; Dimmel & Herbst, 2018; Herbst et al, 2011). The study's goal was to determine whether participants' reactions to the episodes differed depending on the instructional actions that were represented in the different scenarios.

Storyboard Design

Storyboards consisted of 12 – 20 frames of classroom activity that were represented using simple cartoon graphics. They were designed in pairs: Up until a 3-5 frame segment of storyboard (i.e., the *segment of interest*), the storyboards in a pair were identical. The segments of interest for the storyboards were the segments of the storyboard that depicted specific teaching actions that were investigated during the survey experiment. In addition to the segment of interest, each storyboard pair featured a distracter segment where a teacher performed another distinct teaching activity. These activities were considered routine instructional activities, e.g., a teacher asking students to check their work, and they were the same for both the storyboards in a pair. The first and second storyboards targeted how students employed labels and reasons, respectively, when transcribing proofs. The treatment versions of these two storyboards show a teacher departing from the hypothesized norm by interfering with the student's transcription in places where it could be described as lacking mathematical coherence from a disciplinary standpoint—e.g., referring to labeled angles in the written statements of a proof before adding those labels in the diagram (Figure 1). The control versions show a teacher allowing the students to transcribe their proof without interference.

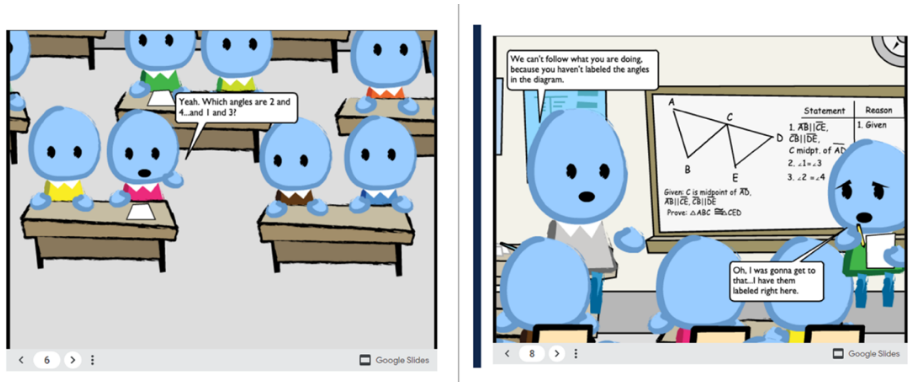

Fig.1 Frames from the labels storyboard

In addition to the replication of the original results, two additional storyboard pairs were designed to investigate how secondary geometry teachers would react to proof presentations that approximated disciplinary communication practices. Storyboard A (Figure 2) depicts an instructional scenario where two students presented a proof at the board. In the treatment version, one student writes the proof and the other verbalizes the proof as it is being written, while in the control version the students took turns transcribing the proof (Figure 2). Storyboard B compared two possible interventions that a teacher could make to steer a student's presentation of a proof toward the multimodal practices that are typical in the discipline. In treatment 1 (Fig. 3, right frame), the teacher asks the student to include appropriate congruence markings and to point at the diagram. In treatment 2, (Figure 3, left frame), the teacher asked the student to provide an overview of the proof strategy and use a conceptual register (Herbst, Kosko, & Dimmel, 2013) to describe the steps in the proof, rather than simply read the statements as they were written on the board. Comments on proof strategy and register switching are touchstones of mathematicians' proof presentations (Weber, 2004).

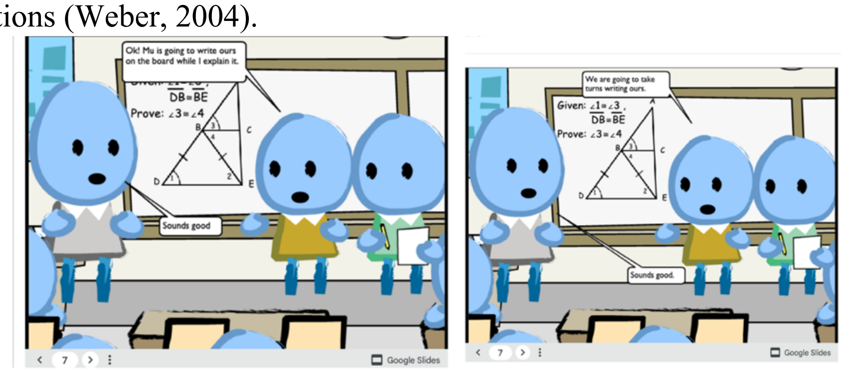

Fig.2 Frames from Alternative Communication Storyboard A

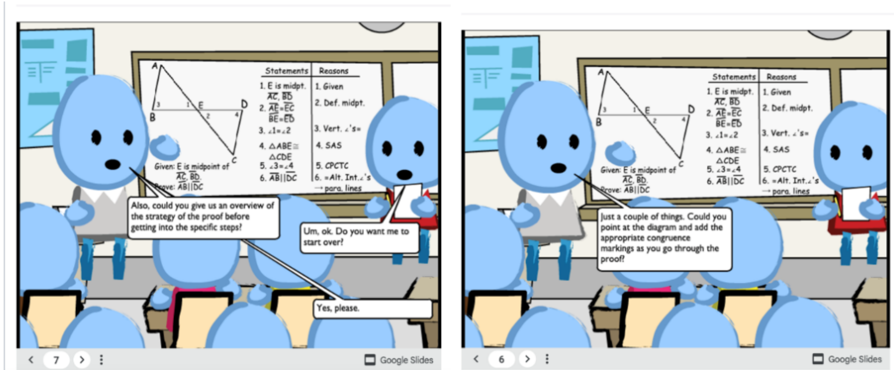

Fig.3 Frames from the Alternative Communication Storyboard B

Participants and Data Collection

The participants for this study were 405 secondary mathematics teachers from 46 states within the United States. All participants were experienced geometry teachers with an average of 7 years of experience teaching geometry. 84.6% of them identified as White, 6.31% as Black, 2.78% as Asian, 2.02% as Hispanic, and 0.76% as other. 60.1% of them were female and 39.9% were male.

Using the storyboards described above as probes, data was collected using a multimedia survey assigned to designated experiment groups. The survey was administered remotely from 2015 – 2016 and was one of a suite of research instruments that were deployed in a large-scale study of the instructional practices of secondary mathematics teachers (Herbst et al., 2015). The experiment groups were designed such that each participant in the group viewed one and only one storyboard in each storyboard pair. Participants were asked the same set of open- and closed-ended questions after viewing each storyboard. For this study, we focused on participants' responses to three closed-ended questions that asked participants to rate, using a 6-valued Likert-like response format, the appropriateness of the teacher's actions. Participants were asked to rate the appropriateness of the teacher's actions across the entire episode and specifically for the actions depicted in the segment of interest and the distracter segment. The rating choices for the closed-ended rating questions were: 1 (very inappropriate), 2 (inappropriate), 3 (somewhat inappropriate), 4 (somewhat appropriate), 5 (appropriate), and 6 (very appropriate).

Analysis

We analyzed participant responses to the closed-ended questions by creating planned comparisons within and between experimental groups. For both the replication storyboards, we generated two sets of hypotheses for the planned comparison: Across condition (three hypotheses, different participants) and within condition (two hypotheses, same participants). For the replication storyboards (Labels and Reasons), across conditions we hypothesized (1) that participants would rate the teacher's work more negatively in episodes that showed a departure from proof transcription than in episodes that showed students transcribing their proofs. We also hypothesized (2) that the mean ratings of *episode appropriateness* (EA) and *segment-of-interest appropriateness* (SIA) would be lower in episodes that depicted a departure from proof transcription compared to the episodes that showed transcription of proofs. Similarly, we hypothesized (3) that there would be no difference in the mean ratings of the *distracter segment appropriateness* (DSA) across conditions (since this action was the same across each storyboard pair). Within conditions, we compared the segment-of-interest and distracter ratings for both the breach and routine conditions. We hypothesized (4) that the mean ratings of SIA would be lower than that of DSA in the treatment condition and (5) that there would be no difference between the ratings in the control condition (since these actions, though different, were both hypothesized to be routine teaching actions with the control conditions). For the alternative communication practice storyboards, we hypothesized that teachers would prefer proof presentations that hewed more closely to the proof transcription norm (the control version of Storyboard A, treatment 2 for Storyboard B). We tested this hypothesis by comparing mean appropriateness ratings across conditions for each of the three appropriateness measures. We conducted a repeated measures ANOVA with planned comparison to test our hypotheses.

Results

Replication Results

For the across conditions comparison, participant ratings of the treatment and control storyboards were as predicted for the episode and segment of interest for the two storyboards that replicated the original study. In comparison to the control storyboards, the treatment storyboards had significantly lower means on the episode appropriateness rating questions (Tables 1 and 2, row 1) and the segment of the interest rating questions (Tables 1and 2, row 2). The distracter segment of the reason storyboard showed a significant difference in mean ratings (Tables 2, row 3). The effect-size statistical analysis showed that while the distracter segments of this storyboard had statistically significant differences in mean ratings, the size of the effect was small, with d =.278. These results thus support the existing literature that when participants view instructional episodes in which teachers allowed students to transcribe their proofs, the work of the teacher were rated higher than in those episodes in which the teacher interfered with the transcriptions.

Table 1: Across condition pairwise comparison for Label Storyboard

Appropriateness Measure	R (n=189)	B (n=212)	μ1-μ2	Std. Error	Sig.	95% confidence Interval for Difference: Lower Bound	Upper Bound
Episode	4.233	3.684	0.549	.125	<.001	.303	.794
Segment of Interest	4.312	2.920	1.392	.143	<.001	1.111	1.674
Distracter Segment	4.741	4.675	0.066	.114	.561	-.158	.290

Based on estimated marginal means
* The mean difference is significant at the .05 level

Table 2: Across condition pairwise comparison for Reason Storyboard

Appropriateness Measure	R (n=189)	B (n=212)	μ1-μ2	Std. Error	Sig.	95% confidence Interval for Difference: Lower Bound	Upper Bound
Episode	3.884	3.052	0.832	.125	<.001	.585	1.078
Segment of Interest	3.984	2.717	1.267	.125	<.001	1.022	1.512
Distracter Segment	4.418	4.090	0.328	.115	.005	.102	.555

Based on estimated marginal means
* The mean difference is significant at the .05 level

For the within condition comparison, the findings shown in Table 4 and 5 provide evidence that the mean ratings of the segment of interest (SIA) were lower than the mean ratings of the distracter segments, for both the Labels and Reasons storyboards. Although this result supports our hypothesis for the treatment groups, it contradicted it for the control groups. The control storyboards depicted routine teaching actions for both the segment of interest and the distracter segments, hence we anticipated that there would be no significant difference between the mean ratings. An effect size analysis for both the Label and Reason storyboards (d = -0.340 and -0.380, respectively) showed that the differences in the means between SIA and DSA were negligible, even if they were statistically significant.

Table 3: Within condition pairwise comparison for Label Storyboard

Condition	SIA	DSA	$\mu 1-\mu 2$	Std. Error	Sig.	95% confidence Interval for Difference		Cohen's D
						Lower Bound	Upper Bound	
Routine	4.312	4.741	-.429	.105	<.001	-.635	-.223	-.340
Breach	2.920	4.675	-1.755	.099	<.001	-1.949	-1.560	-1.328

Based on estimated marginal means
* The mean difference is significant at the .05 level

Table 4: Within condition pairwise comparison for Reason Storyboard

Condition	SIA	DSA	$\mu 1-\mu 2$	Std. Error	Sig.	95% confidence Interval for Difference		Cohen's D
						Lower Bound	Upper Bound	
Routine	3.984	4.418	-.434	.085	<.001	-.601	-.267	-.380
Breach	2.717	4.090	-1.373	.080	<.001	-1.530	-1.215	-1.106

Based on estimated marginal means
* The mean difference is significant at the .05 level

Results: Alternative Communication Practices

In storyboard A, in comparison to the control storyboard, the treatment storyboard had significantly higher means on the episode appropriateness ratings (Table 5, row 1) and the segment of interest ratings (Table 5, row 2). These results were different from what we hypothesized. They suggest that secondary geometry teachers reacted positively to the instructional actions where the teacher enlisted the students to work together to present a multimodal (speaking and writing) version of the proof – we will return to this point in the discussion. The distracter segment showed a significant difference in mean ratings (Tables 5, row 3), however the effect-size analysis revealed that while the distracter segments of this storyboard had statistically significant differences in mean ratings, the size of the effect was small, with d =-0.25.

Table 5: Across condition pairwise comparison for Communication Storyboard A

Appropriateness Measure	R (n=189)	B (n=212)	μ1-μ2	Std. Error	Sig.	95% confidence Interval for Difference Lower Bound	Upper Bound
Episode	4.921	5.321	-.400	.086	<.001	-.570	-.230
Segment of Interest	4.111	4.962	-.851	.116	<.001	-1.079	-.623
Distracter Segment	4.979	5.212	-.233	.091	.011	.054	.413
Based on estimated marginal means							
* The mean difference is significant at the .05 level							

In storyboard B, treatment 1 had significantly higher means on the episode appropriateness ratings (Table 6, row 1) and the segment of interest ratings (Table 6, row 2) in comparison to treatment 2. The distracter segment showed a significant difference in mean ratings (Tables 6, row 3) however the effect-size analysis revealed that the size of the effect was small, with d =.234. Both treatments depicted the teacher asking the students to engage in disciplinary communication practices. For treatment 1, these included gesturing while speaking and marking the diagram. For treatment 2, these included switching from a generic to a conceptual register and offering commentary on the overall plan for the proof. These results suggest that gesturing and marking were seen as more appropriate expectations for teachers to have of student presentations.

Table 6: Across condition pairwise comparison for Communication Storyboard B

Appropriateness Measure	T1 (n=189)	T2 (n=212)	μ1-μ2	Std. Error	Sig.	95% confidence Interval for Difference Lower Bound	Upper Bound
Episode	4.587	3.901	0.686	.119	<.001	.453	.920
Segment of Interest	4.275	3.472	0.803	.144	<.001	.521	1.086
Distracter Segment	5.132	4.910	0.222	.092	.016	.041	.403
Based on estimated marginal means							
* The mean difference is significant at the .05 level							

Discussion and Conclusion

The replication study offered a robust empirical test of the hypothesis that secondary geometry teachers expect student presentations of proofs to default to proof transcriptions (Dimmel & Herbst, 2020). This finding is significant because it provided additional evidence to support an observed social phenomenon. Given the ongoing replication crises in human subject research, such a result is non-trivial (Shrout & Rodgers, 2018). The iteration of the study that investigated teachers' reactions to alternative communication practices is significant because it provided evidence that teachers reacted positively to instructional actions that challenged students to develop their multimodal communication skills. In fact, across all four storyboard pairs, the highest mean ratings for the episode and segment of interest were linked to the storyboard that depicted two students working together to achieve a multimodal presentation of a proof. Also of note, each of the alternatives that were tested in the comparison for Storyboard B had higher mean ratings than the breaches that were depicted in the original study. The implication for teaching is that rather than concentrating solely on teaching students how to write proofs, teachers could provide opportunities for students to learn how to present proofs using approximations to the multimodal practices that are used by experts.

Acknowledgement

This work is supported in part by grants from the Spencer Foundation (second author) and the National Science Foundation (Award #2145517, second author). This research includes secondary data analysis of a data corpus collected with the resources of NSF grant DRL- 0918425 to P. Herbst. All opinions are those of the authors and do not necessarily represent the views of the Spencer Foundation or the National Science Foundation.

References

Artemeva, N., & Fox, J. (2011). The writing's on the board: The global and the local in teaching undergraduate mathematics through chalk talk. Written Communication, 28(4), 345-379.

Ball, D. L., & Bass, H. (2003). Making mathematics reasonable in school. A research companion to principles and standards for school mathematics, 27-44.

Dimmel, J. K., & Herbst, P. G. (2020). Proof transcription in high school geometry: a study of what teachers recognize as normative when students present proofs at the board. Educational Studies in Mathematics, 105(1). https://doi.org/10.1007/s10649-020-09975-y

Dimmel, J. K., & Herbst, P. G. (2018). What details do teachers expect from student proofs? A study of proof checking in geometry. Journal for Research in Mathematics Education, 49(3), 261–291. https://doi.org/10.5951/jresematheduc.49.3.0261

Garfinkel, H. (1963). A conception of and experiments with" trust" as a condition of concerted table actions. The production of reality: Essays and readings on social interaction, 381-392.

Goodenough, W. (1971). Culture, language, and society reading. Massachussets: Addison-Wesley.

Herbst, P., & Brach, C. (2006). Proving and Doing Proofs in High School Geometry Classes: What Is It That Is Going on for Students? (Vol. 24, Issue 1).

Herbst, P., & Chazan, D. (2003). Exploring the Practical Rationality of Mathematics Teaching through Conversations about Videotaped Episodes: The Case of Engaging Students in Proving. In Source: For the Learning of Mathematics (Vol. 23, Issue 1).

Herbst, P., Kosko, K. W., & Dimmel, J. K. (2013). How Are Geometric Proof Problems Presented? Conceptualizing and Measuring Teachers' Recognition of the Diagrammatic Register. North American Chapter of the International Group for the Psychology of Mathematics Education.

Herbst, P., & Chazan, D. (2015). Studying professional knowledge use in practice using multimedia scenarios delivered online. International Journal of Research and Method in Education, 38(3), 272–287. https://doi.org/10.1080/1743727X.2015.1025742

Herbst, P., Chazan, D., Chen, C. L., Chieu, V. M., & Weiss, M. (2011). Using comics-based representations of teaching, and technology, to bring practice to teacher education courses. ZDM, 43, 91-103.

Herbst, P. G. (2006). Teaching Geometry with Problems: Negotiating Instructional Situations and Mathematical Tasks. In Source: Journal for Research in Mathematics Education (Vol. 37, Issue 4).

Kokushkin, V., Norton, A., Johnson, E., Alibali, M., Doolittle, P., Elgart, A., & Arnold, R. (2022). The Role of Students' Gestures in Offloading Cognitive Demands on Working Memory in Proving Activities.

Kosko, K. W., & Wilkins, J. L. (2010). Mathematical communication and its relation to the frequency of manipulative use. International Electronic Journal of Mathematics Education, 5(2), 79-90.

Lai, Y., & Weber, K. (2014). Factors mathematicians profess to consider when presenting pedagogical proofs. Educational Studies in Mathematics, 85, 93-108.

Lee, C. (2006). Language for learning mathematics: assessment for learning in practice: Assessment for learning in practice. McGraw-Hill Education (UK).

National Council of Teachers of Mathematics. (2002). Principles and standards for school mathematics. Reston, VA: Author.

Shrout, P. E., & Rodgers, J. L. (2018). Psychology, Science, and Knowledge Construction: Broadening Perspectives from the Replication Crisis. Annual review of psychology, 69, 487–510. https://doi.org/10.1146/annurev-psych-122216-011845

Stylianides, A. J. (2019). Secondary students' proof constructions in mathematics: The role of written versus oral mode of argument representation. Review of Education, 7(1), 156–182. https://doi.org/10.1002/rev3.3157

Weber, K. (2004). Traditional instruction in advanced mathematics courses: A case study of one professor's lectures and proofs in an introductory real analysis course. Journal of Mathematical Behavior, 23(2), 115–133. https://doi.org/10.1016/j.jmathb.2004.03.001