



Proof and proving in school and university mathematics education research: a systematic review

Gabriel J. Stylianides¹ · Andreas J. Stylianides² · Andreas Moutsios-Rentzos³

Accepted: 10 August 2023 / Published online: 30 August 2023
© The Author(s) 2023

Abstract

This systematic review aims to provide a complementary to existing synopses of the state-of-the-art of mathematics education research on *proof* and *proving* in both school and university mathematics. As an organizing framework, we used Cohen et al.'s triadic conceptualization of instruction, which draws attention not only to the main actors of the didactical process (i.e., the *Teacher* and *Students*) and the *Content* around which the actors' work is organized (herein, content related to proof and proving), but also to the relationships among the actors and the content. Out of the 103 papers we reviewed, almost half fell in the Student–Content category, which is consistent with the existence of a substantial number of frameworks, methods, and research findings related to students' engagement with proof and proving. About a quarter of the papers fell in the Student–Teacher–Content category, which reflects an emphasis on viewing instructional practice in proof and proving in a holistic, systemic way. Only few papers fell in the categories that did not include Content in them, namely, the categories of Student, Teacher, and Student–Teacher; this suggests mathematics education research on proof and proving has a strong disciplinary identity, which potentially differentiates it from other mathematics education research strands. About a fifth of the papers were oriented towards 'breaking ground' through making an explicit theoretical and/or methodological contribution (Student–Teacher–Content and Content were the main categories where such contributions appeared), whilst the majority of the papers were focused on 'building ground' through elaborating or employing existing methodological and/or theoretical approaches.

Keywords Proof · Proving · Instruction · School · Teacher education · University · Systematic review

1 Introduction

The concept of *proof* and the respective activity of *proving* have received growing attention by mathematics education researchers in the past few decades, and this attention is reflected in the upsurge of publications on didactical, epistemological, cognitive, and other aspects of proof, proving, and related concepts such as *argumentation* and *reasoning-and-proving*. The high level of research attention to proof and proving, which has spanned all levels of schooling

(including the elementary) and relevant areas of research at the university level (including teacher education), has been justified from multiple perspectives. For example, from a philosophical/curricular perspective, the pivotal role of proof in disciplinary experts' work—to the extent that proof has been called 'the soul of mathematics' (Schoenfeld, 2009)—necessitates that mathematics curricula that have integrity should offer to students authentic encounters with proof from the beginning of their mathematical education (Bruner, 1960; A. Stylianides, 2016). Relatedly, from a pedagogical and epistemological perspective, mathematically and cognitively appropriate learning opportunities for students to engage with proof can afford students with a basis for deep learning: assertions and new knowledge get accepted because they make sense mathematically rather than based on the authority of the teacher or the textbook (Bass & Bass, 2003).

Unlike prior reviews of mathematics education research in this area that were conducted by us and others over the

✉ Gabriel J. Stylianides
gabriel.stylianides@education.ox.ac.uk

¹ Department of Education, University of Oxford, Oxford, UK

² Faculty of Education, University of Cambridge, Cambridge, UK

³ Department of Pedagogy and Primary Education, National and Kapodistrian University of Athens, Athens, Greece

past couple of decades (Harel & Sowder, 2007; Mariotti, 2006; Mariotti et al., 2018; Sommerhoff & Ufer, 2019; Stylianides et al., 2016, 2017, 2022), this one is not guided by an explicit proof-related theoretical framework or by pre-identified proof-related themes emerging from prior research. Rather, as we explain in the following section, we use as an organizing framework for the review Cohen et al.'s (2003) triadic conceptualization of instruction. This choice reflects our attempt to address a gap in the existing reviews of mathematics education research on proof and proving, namely, the lack of a bird's eye view of the position of this body of research relative to the various elements of the instructional triangle (student–teacher–content), which has attracted considerable attention by educational researchers internationally and across many mathematics education topics (Goodchild & Sriraman, 2012). By utilizing this framework to organize our review, we explicitly investigate whether and to what extent the didactical/instructional relationship is addressed in the recent body of research that focused on proof and proving.

To conclude, with this review, we attempt to offer an instruction-focused synopsis of the current state-of-the-art of mathematics education research on proof and proving at both the school and university levels and to identify current and emerging trends in the field by addressing the following research question: *How are recent mathematics education studies on proof and proving mapped across Cohen et al.'s (2003) triadic conceptualization of instruction, and what themes emerge from this mapping?*

2 Elaboration on the scope of the review and review procedure

In three recent reviews of research on proof and proving published as handbook chapters, we documented the state-of-the-art in this area from complementary angles. In Stylianides et al. (2016), we followed a systematic approach to identify themes amongst relevant papers focusing on proof and argumentation and that were published in the proceedings of the International Group for the Psychology of Mathematics Education (PME), covering the period 2005–2015. Before us, Mariotti (2006) had reviewed themes emerging from proof-related papers published in the PME proceedings covering an earlier period, while Mariotti et al. (2018) did the same but for proof-related papers published in the proceedings of the Congress of European Research in Mathematics Education (CERME) over a twenty-year period. Sommerhoff and Ufer's (2019) review also focused on PME proceedings, this time covering the period 2010–2014, with particular attention to argumentation and proof studies in secondary and tertiary education. Sommerhoff and Ufer used pre-determined categories, derived from prior research, to

identify the extent to which different prerequisites and goals of argumentation and proving processes were investigated within PME reports in their sample.

In Stylianides et al. (2017), we conducted a narrative review of proof-related research, published mostly in refereed journals prior to 2016, from the perspectives of proving as problem solving, proving as convincing, and proving as a socially embedded activity. Before us, Harel and Sowder (2007) had conducted another narrative review covering a similarly broad body of published research related to proof but using their notion of 'proof schemes' as a theoretical framework.

Finally, in Stylianides et al. (2022), we conducted yet another narrative review, which was not restricted to a specific time-period, in order to define and exemplify authentic classroom mathematical activity in the area of proving. The conclusions and directions for future research of the current and our previous reviews were determined by each review's particular focus or theoretical framework and period of interest. For complementarity, and for the reason we explained previously to address a gap in the existing reviews of mathematics education research on proof and proving, we defined the scope of the current review differently.

Specifically, we offer a synopsis of the state-of-the-art in the area of proof and proving through a review of relevant papers in English that were published (including 'online first') between 1/1/2018 and 1/6/2022 in the mathematics education research journals in Table 1, using as an organizing framework Cohen et al.'s (2003) notion of *instructional triangle*. We restricted our review to the journals that were graded as A*, A, or B by the expert panel convened by the Education Committee of the European Mathematical Society (EMS) and the Executive Committee of the European Society for Research in Mathematics Education (ERME)

Table 1 The journals included in the review and respective grade based on the grading reported in Toerner and Arzarello (2012, p. 53)

Grade	Journal title
A*	Educational Studies in Mathematics Journal for Research in Mathematics Education
A	For the Learning of Mathematics Journal of Mathematical Behavior (The) Journal of Mathematics Teacher Education Mathematical Thinking and Learning ZDM—Mathematics Education
B	International Journal of Mathematics Education in Science and Technology International Journal of Science and Mathematics Education Mathematics Education Research Journal Recherches ed Didactique des Mathématiques ^a Research in Mathematics Education

^aOnly papers in English were included in this review

(Toerner & Arzarello, 2012). In making the decision to restrict our review to those journals, we were guided by a need for a systematic and practical procedure rather than elitism. We acknowledge that important research on proof and proving is published also in general education research journals, in conference proceedings, in books and book chapters, and in other languages that we did not consider in the review. Indeed, some of our own (including non-English) research has been published in outlets that we reluctantly excluded from this particular review but, as we explained previously, we addressed in other reviews.

While we explained in detail elsewhere our own perspective on the meanings of proof and proving (e.g., Stylianides, 2007; Stylianides et al., 2017), for the purposes of this review we followed an inclusive approach and considered all papers that we obtained from searching the journals in Table 1 and that included *proof*, *proving*, or *argumentation* in at least one of their title/abstract/keywords. To organize the review, we used Cohen et al.'s (2003) triadic conceptualization of instruction according to which instruction “consists of interactions among teachers and students around content, in environments” (p. 122). This conceptualization draws attention not only to the main actors of the teaching and learning process—the *teacher* (T) and the *students* (S)—and the *content* (C) around which the actors' work is organized (in this case, content related to proof and proving in mathematics), but also to the relationships among the actors and the content.

There is no suggestion that the merits of a paper are dependent on whether it addresses a single vertex of the instructional triangle (T, S, C) or the relationships between two or all three vertices (SC, ST, TC, STC). However, it is posited that it is important for us as a field to get a sense of the relative attention that mathematics education researchers have paid to the various elements of instruction in the area of proof and proving, and whether there are any differences between levels of education.

Furthermore, there is no suggestion that the use of the (particular version of the) instructional triangle framework is unproblematic. We share Goodchild and Sriraman's (2012) view that, “[i]n many ways the ‘simple’ representation of didactical systems depicted in the didactic [instructional] triangle is argued to be inadequate” (p. 584), and indeed these inadequacies have motivated several extensions of and modifications to the framework to accommodate various research aims or theoretical perspectives (some of these extensions or modifications are discussed in the papers of the special issue edited by Goodchild and Sriraman, 2012). Yet, as Goodchild and Sriraman observed further, “all papers [in their special issue] confirm the central position of mathematics, learner and teacher in researching and theorizing teaching–learning processes in mathematics classrooms” (p. 584). On this basis, we judged that the ‘simple’

(in Goodchild & Sriraman's, 2012 terms) version of the instructional triangle was the most appropriate for a review of a broad body of research, such as ours, that considers a multitude of papers that do not share a common theoretical perspective or research aim.

In conducting the analysis, we first categorized each of the papers we identified as relevant for this review according to whether the paper's (main) *focus* was on S, T, C, SC, ST, TC, or STC. Similarly to Sommerhoff and Ufer's (2019) review, the notion of ‘focus’ is important to understand our coding process. For example, consider a paper investigating students' views of argumentation or proof in geometry in the context of classroom discussions. Building on the same example, if the teacher, although present in those discussions and potentially influential of the issues discussed by the students, did not feature prominently in the analysis or the research questions addressed in the paper, the paper would be coded as SC. Similarly, the paper would be coded as TC in case the paper focused on the teachers' views while the students' role was kept in the background of the reported research; this would still be the case even if the discussions could not have happened without the students' involvement. If the teacher's role and the students' role were both explicitly investigated in the paper, then the paper would be coded as STC. In case that the students' and the teacher's presence were both missing or were not at the crux of the research questions of the paper, focusing only on aspects of argumentation or proof in geometry, then the paper would be coded as C. A paper that investigated teacher–student interactions, with the mathematical content (e.g., argumentation and proof in geometry) appearing not to play a particular role in the reported research, would be coded as ST. Accordingly, papers focusing only on the students or the teacher would be coded respectively as S or T.

The focus of a paper was also further identified in accordance with the educational system of reference. Papers that examined (preservice) teachers as learners were categorized as S, as their role in the respective educational system (i.e., university) is that of a student, whereas those that examined them as teachers of mathematics were categorized as T, as their role in the respective educational system (i.e., school) is that of a teacher. Similarly, papers that examined mathematicians from their roles as instructors in university were categorized as T. Also, we recorded the educational level of each paper: *school*, *university*, *both* (school and university), or *discipline* (in case the paper was not concerned with a particular level but rather was about the discipline in general). Teacher education papers with preservice teachers were classified as university level.

As expected, all papers made some sort of theoretical and/or methodological *contribution*. Some papers were more oriented towards (focused on) *building ground*, that is, they mostly contributed to research knowledge by further

developing or implementing existing theoretical ideas, methods, or findings. Other papers were more oriented towards (focused on) *breaking ground*, that is, they were more explicitly written with the aim of developing or implementing new theoretical ideas or methods. We further coded the ‘breaking ground papers’ according to whether they explicitly made a *theoretical* and/or a *methodological* contribution (cf. Hanna & Knipping, 2020). We clarify that, for papers coming from the same group of authors making the same kind of methodological or theoretical contribution, only the first chronologically publication in our list was recorded as breaking ground; we viewed follow-up publications as strengthening or elaborating on that contribution and thus recorded them as building ground.

Following the journal search as we described earlier, we obtained an initial set of 126 potentially relevant papers. After screening the papers’ abstracts and, frequently, their full version as well, we excluded papers that were book reviews, science focused, or corrections/comments on other papers. This left us with a final set of 103 papers that we provide in Supplementary Materials. The coding procedure was as follows: two of us considered all categorizations separately and, subsequently, we resolved all disagreements through discussion (although inter-rater agreement was not a major concern for this review, our initial agreement was ‘good’ to ‘excellent,’ $Kappa=0.718$). To meet the restrictions of how many of the 103 papers we could reference in this review, as well as which subset of these papers we would annotate in the bibliography of the review, we aimed for a

representation of the themes that emerged from the review, but also a representation of authors and some of the trends that emerged recently during key conferences in the field (not only CERME and PME, but also of the International Congresses on Mathematical Education [ICME]). The identification of themes and the selection of papers we used to illustrate them, inevitably, was influenced by our subjective judgment but, at the same time, was guided by our knowledge of the field and the findings of our review.

3 Findings and discussion

Tables 2 and 3 summarize the findings of our analysis of the 103 papers. Specifically, the tables show, respectively, the distribution of these papers by level (school, university, school and university, or discipline) and contribution (breaking ground or building ground) across the seven possible combinations of the instructional triangle categories (S, T, C, SC, ST, TC, STC). There are several broad observations to draw from the tables:

- 1) Almost half of the papers in the sample (49 out of 103) fell in the SC category. This finding accords with the observation that the field thus far has produced a substantial number of frameworks and research findings related to students’ engagement with proving, especially students’ difficulties with proving tasks (Mariotti et al., 2018; Stylianides et al., 2017).

Table 2 Distribution of papers in the sample (N= 103) by level according to their instructional triangle focus

Instructional Triangle Focus		Level				
		School	University	School & University	Discipline	Total
Student (S)	<i>f</i>	1	3	0	0	4
	%	2.7% ^a	5.7%	0.0%	0.0%	3.9%
Teacher (T)	<i>f</i>	0	0	0	0	0
	%	0.0%	0.0%	0.0%	0.0%	0.0%
Content (C)	<i>f</i>	2	6	1	4	13
	%	5.4%	11.3%	11.1%	100.0%	12.6%
Student–Content (SC)	<i>f</i>	17	28	4	0	49
	%	45.9%	52.8%	44.4%	0.0%	47.6%
Student–Teacher (ST)	<i>f</i>	0	1	0	0	1
	%	0.0%	1.9%	0.0%	0.0%	1.0%
Teacher–Content (TC)	<i>f</i>	9	3	1	0	13
	%	24.3%	5.7%	11.1%	0.0%	12.6%
Student–Teacher–Content (STC)	<i>f</i>	8	12	3	0	23
	%	21.6%	22.6%	33.3%	0.0%	22.3%
Total	<i>f</i>	37	53	9	4	103
	%	35.9% ^b	51.5%	8.7%	3.9%	100.0%

^aWithin each category of Level

^bWithin Level

Table 3 Distribution of papers in the sample (N= 103) by contribution (breaking ground or building ground) according to their instructional triangle focus

Instructional Triangle Focus	Contribution						
	Breaking Ground				Building Ground	Grand Total	
	Theoretical	Methodological	Theor. & Method.	Total			
Student (S)	<i>f</i>	0	1	0	1	3	4
	%	0.0% ^a	14.3%	0.0%	5.0%	3.6% ^c	3.9% ^d
Teacher (T)	<i>f</i>	0	0	0	0	0	0
	%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Content (C)	<i>f</i>	6	2	0	8	5	13
	%	60.0%	28.6%	0.0%	40.0%	6.0%	12.6%
Student–Content (SC)	<i>f</i>	3	2	0	5	44	49
	%	30.0%	28.6%	0.0%	25.0%	53.0%	47.6%
Student–Teacher (ST)	<i>f</i>	0	1	0	1	0	1
	%	0.0%	14.3%	0.0%	5.0%	0.0%	1.0%
Teacher–Content (TC)	<i>f</i>	0	0	0	0	13	13
	%	0.0%	0.0%	0.0%	0.0%	15.7%	12.6%
Student–Teacher–Content (STC)	<i>f</i>	1	1	3	5	18	23
	%	10.0%	14.3%	100.0%	25.0%	21.7%	22.3%
Total	<i>f</i>	10	7	3	20	83	103
	%	50.0% ^b	35.0%	15.0%	100.0%	100.0%	100.0%

^aWithin each category of Breaking Ground Contribution

^bWithin Breaking Ground Contribution

^cWithin Building Ground Contribution

^dWithin Contribution

- 2) About a quarter of the papers in the sample (23 out of 103) fell in the STC category. This reflects a current emphasis in the literature (Mariotti et al., 2018) to viewing instructional practice in the area of proof and proving in a rather holistic way, considering all actors involved (students and teacher) as well as their interactions with the content. Although our review did not cover the pre-2018 period, we believe this is an emerging trend in the literature partly as a response to calls for more classroom-based studies (Mariotti, 2006; Mariotti et al., 2018) and research on classroom-based interventions (G. Stylianides & A. Stylianides, 2017; G. Stylianides et al., 2017) to better understand the teacher’s role in facilitating students’ interactions with proving (even that of teachers in the shadow education system; Moutsios-Rentzos & Plyta, 2019). Moreover, this trend echoes related efforts in the broader field of mathematics education that address emerging complex mathematics teaching and learning environments (Moutsios-Rentzos et al, 2017).
- 3) There were only few papers (ranging from 0 to 4) in the categories that did not include C in them, namely, the S, T, and ST categories. This shows that, recent research in the area of proof and proving, rarely consid-

- ered students or teachers in isolation of the mathematical content within which proof and proving occurs. This is especially true in the case of the teachers who seem to be a focus of investigations predominantly when content is involved (never on their own and only once with students). It suggests further that mathematics education research on proof and proving has a strong disciplinary identity (Mariotti et al., 2018), which potentially differentiates it from other mathematics education research strands that treat (for good reasons) the content more as a research context rather than as a core part of it.
- 4) More papers focused on the university level as compared to the school level (53 vs. 37, respectively), with 9 papers pertaining to both. The remaining 4 papers related to the discipline and were all in the C category. Although the higher representation of the university papers in the sample may relate to our decision to allocate to this level teacher education papers with preservice teachers, it is also indicative of the broad educational focus of the researchers in the field as reflected, for example, in the contributions in the proof-related working groups of the recent ICME and CERME congresses.

- 5) Twenty out of the 103 papers' contribution was deemed to be breaking ground, with 10 of those papers making an explicit theoretical contribution, 7 a methodological contribution, and 3 both. The representation of these 20 papers was uneven across the seven instructional triangle categories. More than half of the papers in the C category made such a contribution (8 out of 13); about one in five in the STC category (5 out of 23); and about one in ten in the SC category (5 out of 49). The remaining 2 papers which made such a contribution belonged to the S and ST categories.

The above key findings are diagrammatically summarized in Fig. 1, where the instructional triangle is used as both an organizational and a communicational scheme of the perspective employed in this review and its results.

Next, we discuss and illustrate the four instructional triangle categories with the most papers in them (SC, STC, C, and TC) to deepen understanding of the kind of research that has been conducted in relation to each of them. In our discussion we consider, as appropriate, papers that we deemed to make a breaking ground contribution. The discussion of the papers of each category is organized through themes that we inductively (and to some extent, subjectively) identified and, subsequently, conceptually linked with the broader mathematics education research including research in the field's major conferences. Given that all four categories

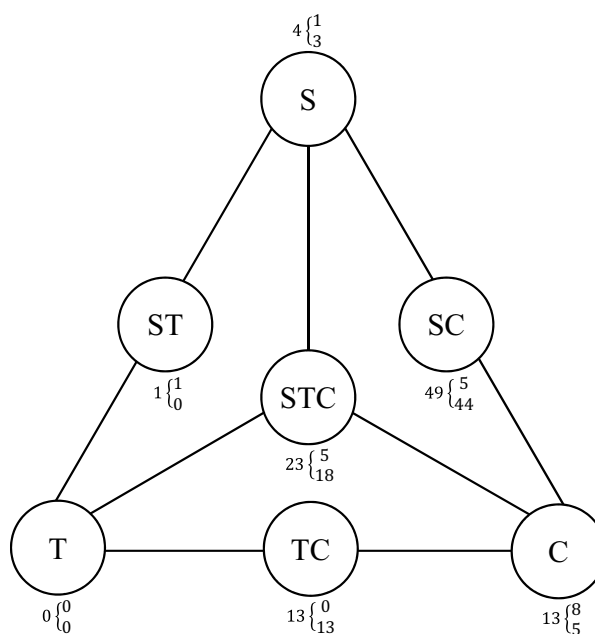
include C in them, we begin with the C category and move next to the SC and TC categories before we discuss the STC category last.

3.1 Content

Almost half of the 13 papers that addressed primarily issues of *proof content*, and were thus included in this category, focused on the university level or were about the discipline more generally. A couple of these papers discussed the role of computer-based tools, such as digital proof assistants (Hanna & Yan, 2021), for facilitating proof construction (see, also, Hanna et al., 2019). While the main concern in these papers was on the affordances of the technology, especially the ways it can mediate access to proof content, implications for mathematics instruction were interwoven with the discussion as a justification for the need for researchers to further explore these affordances.

A few other papers explored *theoretical aspects of proof*, such as its explanatory and convincing functions (Lockwood et al., 2020), its relationship with the notions of evidence and derivation (Aberdein, 2019), and its association (tension) with the notions of freedom and enforcement (Nickel, 2019). These papers made breaking ground contributions by deepening the field's understanding of existing theoretical constructs relevant to proof and their interrelationships. Implications for mathematics instruction are more immediate in some of them than in others. Relatedly, Mejía-Ramos

Fig. 1 A diagrammatic summary of the findings of this review



Notes

Papers were mapped according to their main focus on the label of the respective circle (S, T, C stand for Student, Teacher, Content, respectively)

f_{bugr}^{brgr} , f is the total number of papers assigned to a circle, with *brgr* referring to the number of 'breaking ground' contributions, and *bugr* to 'building ground' contributions

and Weber's (2020) paper, which we deemed to have also made a (methodological) contribution, problematized the methods researchers use to gain insights into mathematical practice and called for exercising caution in how such insights ultimately inform mathematics instruction. The relationship between disciplinary and school mathematics practice, particularly the conditions under which the latter may be characterized as "authentic" in its reflection of the former, was further discussed in A. Stylianides et al. (2022) (not in the sample of 13 papers in this category).

Finally, only two of the papers in this category related exclusively to *school mathematics*, and both explored cultural dimensions of the place of proof in the secondary school curriculum. One of them reported a cross-country comparative textbook analysis (Bergwall, 2021) extending previous textbook-analysis work in the broader area of proof reviewed in G. Stylianides et al. (2017) and called for in Mariotti et al. (2018). The other paper, by Shinno et al. (2018), which we deemed to have also made a breaking ground (methodological) contribution, elaborated an epistemological model for understanding what constitutes proof in the curricula of different countries.

3.2 Student-Content

As can be seen in Table 2, the SC papers spanned educational levels, with 28 papers focusing on the university level, 17 on the school level, and 4 papers on both school and university. Yet some common themes emerge when looking across all 49 papers with a student-content focus.

A main theme relates to papers focusing on the role and use of *examples* in proving, including the use of *counter-examples* (see also Boero, 2022). Several of these papers derived from a special issue in *The Journal of Mathematical Behavior* (Knuth et al., 2019) that addressed students' use of examples in proving and, in particular, ways in which example-based reasoning can be productive for students in the proving process. Knuth et al. (2019) discussed example-based reasoning as an important object of study and explored how examples can play a foundational role not only in the development and exploration of conjectures, but also in the development of proofs for the conjectures. Aricha-Metzer and Zaslavsky (2019) conducted individual interviews with secondary and undergraduate students (majoring in mathematics or mathematics-related subjects) on tasks that called for conjecturing and proving. They found a relatively strong tendency among students to use examples generically, which they associated with a productive use of examples in terms of developing a proof or a sound justification that can lead to a proof. Other studies involved an exploration of how in-service teachers interpret, understand, and use generic examples in their proving activities (Dogan & Williams-Pierce, 2021), a comparison between mathematicians and students'

example use (Lynch & Lockwood, 2019), and an exploration of structural aspects of generic examples to understand what makes them potentially opaque for learners in the domain of multiplication (Rø & Arnesen, 2020). Lew and Zazkis (2019) examined undergraduate mathematics students' at-home work on prove-or-disprove tasks to better understand the interactions between example/counterexample generation activities as students try to prove the truth or falsity of mathematical claims.

Another main theme relates to papers that investigated participants' *understandings and conceptions of proof*. Brown (2018) and Antonini (2019) explored issues of conviction in the context of indirect proofs. Brown employed a comparative selection of tasks to investigate undergraduate mathematics students' views about the convincingness of indirect versus direct proofs. She found that the form of indirect proof (contraposition or contradiction) may make a difference for students and that students' views may be more nuanced than previously considered. Antonini explored the issue of the intuitive acceptance of proof by contradiction with secondary and university students in the domain of geometry. He observed, through analysis of task-based interviews, that students can produce indirect argumentation (by starting with the assumption that a claim is false) as a compromise between proof by contradiction and the need for a more evident argument. Other studies in this area included Weber et al.'s (2022) investigation of how mathematicians use proofs to obtain conviction and certainty, Stewart and Thomas's (2019) examination of undergraduate students' perspectives on proof in linear algebra, Campbell et al.'s (2021) comparison of secondary students' written and oral representations of mathematical arguments to investigate whether the two modalities portray similar understandings, Yopp et al.'s (2020) investigation of secondary students' conceptions about the validity of a direct argument after they participated in instruction that addressed issues of constructing and critiquing arguments for a general claim, and Blanton et al.'s (2022) examination of how young learners (Grades K-1) come to construct viable mathematical arguments to justify generalizations about even and odd numbers. Finally, Davies et al. (2020) used comparative judgement to investigate undergraduate students' proof comprehension based on students' summaries of a given proof. This is one of the papers we identified as making also a breaking ground (methodological) contribution: it shows that comparative judgement can help produce valid and reliable assessments of the quality of students' proof summaries, thereby offering a new way to assess proof comprehension.

Few studies in this category explored student-content issues in *technological environments*. One of these studies was Fujita et al.'s (2019) exploration of how learners can be supported to overcome logical circularity during their proof construction process in the context of a web-based

proof learning support system. Their data derived from three strategically selected cases involving secondary students and prospective elementary school teachers. Although the computer-based environment was enough for some students to progress in their thinking, for others teacher intervention was necessary for them to begin to realize when their proof fell into logical circularity. In a different study, Cirillo and Hummer (2021) used Smartpen technology (i.e., Livescribe pens) to audio-record secondary school students' explanations of their thinking and capture their pen strokes as they worked through geometry proof problems. Also, some of the students in Antonini's (2019) study we mentioned earlier worked in a Dynamic Geometry Environment.

A handful of studies examined students' *proof writing or reading* of proof scripts. Azrou and Khelladi (2019) conducted a case study of undergraduate students' proof writing to understand why students' efforts to write a proof can result in a disorganized, unclear draft. Dawkins and Zazkis (2021) used moment-by-moment think-aloud protocols to understand novice and experienced undergraduate students' processes of reading mathematical proof. Finally, Ahmadpour et al. (2019), in a study that we identified to also make a breaking ground (theoretical) contribution, present a model for describing the growth of students' understandings when reading a proof; the model comprises two paths relating to the semantic and syntactic levels.

A small number of studies in this category involved analyses of prospective teachers' *proof constructions and evaluations* (e.g., Ko & Rose, 2018; Yee et al., 2018). Finally, we found very few studies in the student-content category focusing on students' *ffective aspects*, which accords with the observation that few studies in the field focus on these aspects (e.g., Moutsios-Rentzos & Kalozoumi-Paizi, 2017). One exemption was Ayalon et al.'s (2022) in-depth qualitative study that considered how secondary students' emotions during an argumentative discourse relate to their learning of real-life functional situations.

3.3 Teacher-Content

The papers in this category addressed the interaction between teacher and proof content, and they tended to focus on the secondary school level in the domain of geometry. A smaller number of papers viewed (research) mathematicians in the role of the teacher and were thus classified under the university level. In almost all the papers, student learning or perceptions were also considered, but we deemed these to be more part of the context and less an object of study. The focus in this category is consistent with Mariotti et al.'s (2018) observation of a relatively recent attention by CERME papers on the teacher (versus the student) and their call for more studies on specific teacher competencies

required for designing and managing didactic situations concerning argumentation and proof.

Some school-level papers explored teachers' proof-related *content knowledge or perspectives on proof-related content*. For example, Karpuz and Atasoy (2020) explored Turkish secondary-school mathematics teachers' content knowledge of the logical structure of proof in geometry; they reported teachers' challenges about how to link information from geometrical figures to established knowledge, or how to avoid sweeping generalizations based on geometrical figures. In another study at the secondary school level in geometry, Aaron and Herbst (2019) examined American high-school teachers' perspectives on the interplay between conjecturing and proving; they found that teachers favored the separation of these two activities, perceiving them as having different goals and requiring different resources. Considering these findings against those of another study that examined how a research mathematician conjectured and proved (Fernández-León et al., 2021), one appreciates the need for further enhancement of the teacher-content relationship at the school level through, for example, targeted professional development (e.g., Kazemi et al., 2021).

Another area of attention was teachers' *interactions with proof-related tasks* (Ayalon & Hershkowitz, 2018; Baldinger & Lai, 2019; Rogers & Kosko, 2019). For example, Ayalon and Hershkowitz (2018) examined Israeli secondary school teachers' choices of textbook tasks that they thought had potential to encourage proof-related activity in the classroom. Rogers and Kosko (2019) examined similar issues but from the perspective of teachers as task designers and with a scope that spanned the school and university levels: they asked their participants—elementary school teachers and university mathematics instructors—to design tasks that could support students' ability to create and critique mathematical arguments. Teachers' task choices or task design decisions offered useful insights into the knowledge, values, and dispositions underpinning their proof-related instruction.

3.4 Student-Teacher-Content

The papers in this category are characterized by their attempts to provide a more holistic approach to teaching and learning interactions by including both the main educational protagonists (teachers and students) and the respective content. Moreover, in STC, we identified the highest ratio of papers classified as making a breaking ground contribution, while the three papers that were considered to explicitly make both theoretical and methodological contributions were all in this category; this is in line with our aforementioned hypothesis that the STC category signifies an emerging trend in the field and is consistent with relevant calls for more research in this area (Stylianides & Stylianides, 2017;

Stylianides et al., 2017). The STC papers may be grouped into four themes, according to which the rest of this section is structured.

First, almost half of the papers in the STC category discussed *teaching approaches with respect to proof*, including at the school or university levels. Some papers focused on the teacher's role in supporting the students' learning at school (e.g., Komatsu & Jones, 2022; Zhuang & Conner, 2022), while others investigated ways of supporting preservice (secondary or primary) teachers' learning or teaching skills (e.g., Buchbinder & McCrone, 2020; Zambak & Magiera, 2020). A few papers concentrated on university teaching practices about proof (e.g., independent proof reading; Pinto & Karsenty, 2018). Stylianides and Stylianides (2022) is an example of a paper that purposefully considered both school and university by discussing a coherent approach of introducing proof to secondary students and preservice elementary teachers. Their argument for implementing a coherent approach to introducing proof across the different educational protagonists adds to the field's theoretical perspective about the teaching of proof.

Second, some papers concentrated on *proof norms and criteria*, employing diverse perspectives. For example, Dimmel and Herbst (2020) utilized storyboard representations of instructional situations to investigate the teachers' expectations of their students presenting their work at the board. Other researchers employed a broad perspective to explore the potential convergences and divergences in proof norms (e.g., by including lectures, textbooks, home assignments, and feedback; Pinto & Karsenty, 2020) or in proof criteria (e.g., by including school students, university students, and mathematicians; Sommerhoff & Ufer, 2019).

Third, few papers appeared to focus on explicitly investigating *novel theoretical frameworks and/or methodological approaches* to challenge current perspectives, thereby being coded as making breaking ground contributions. One of those papers was by Gabel and Dreyfus (2020) who discussed the links between rhetoric and proving in the context of the university teaching of proof, thus making a theoretical and methodological contribution to the field about the notion of proof and proving. In addition, Tabach et al. (2020) methodologically and theoretically contributed by networking two approaches to link the mathematical progress of individuals, small groups, and the whole class.

Finally, the *linguistic and the discursive aspects* of proof and proving appeared to also attract the interest of various researchers. For example, Kontorovich (2021) employed a commognitive framework to investigate a topologist's feedback on her students' proofs, thus methodologically contributing to the field. This interest is also evident in recent conferences (e.g., Boero & Turiano, 2023; Moutsios-Rentzos, 2022; a joint session of the argumentation and proof and the language thematic working groups in CERME 12).

4 Concluding remarks

In this review, we aimed to provide a complementary to existing synopses of the current state-of-the-art of mathematics education research on proof and proving by using a triadic framework of instruction (Cohen et al., 2003) to conceptually and methodologically organize our review. We also identified the educational level of the papers included in the review, showing that, at least recently, the field appears to be relatively balanced between the university level (which includes teacher education) and the school level. Before we further discuss the findings from our review, we acknowledge limitations of our approach.

Although our decision to use the 'simple' version of the instructional triangle framework was deemed appropriate for the scope of this review (to provide a bird's eye, instructionally focused view of this area of research), in line with Goodchild and Sriraman's (2012) commentary on the instructional triangle, we acknowledge that using other, more elaborated versions of the framework might have illuminated different aspects of the papers we reviewed. The fact that we did not consider papers published in conference proceedings is another limitation even though, as we mentioned previously, some past reviews focused exclusively on conference proceedings (Mariotti, 2006; Mariotti et al., 2018; Sommerhoff & Ufer, 2019; Stylianides et al., 2016). We acknowledge further that the picture we obtained from our review might have differed had we been more inclusive not only in terms of conference proceedings, but also in terms of the journals and the publication period we covered in the review. To mitigate the effects of these limitations, we considered broader literature in our discussion including papers (published in conference proceedings and other outlets) that did not meet the strict parameters of our search criteria.

Notwithstanding these limitations, we argue that the employed approach served its purpose by revealing specific trends in the recent proof and proving literature with respect to instruction. First, the role of Content (C) is central and multifaceted to the recently published work in this field, as almost all the papers (98/103) included C in their main foci including a good number of papers (13/103) that included C as their only focus. At the same time, the papers focusing only on C appeared to constitute a notable category in the sense that they included the highest ratio of papers making a breaking ground contribution to the field. Moreover, C was significant enough for some papers to transcend educational levels and to concentrate on the discipline itself (the papers focused on the discipline were all assigned to C). The significant role of C in our review was not unexpected as the notion of proof has a strong disciplinary identity (Mariotti et al., 2018); proof is at the

crux of the ontology and epistemology of mathematics (e.g., Ernest, 2018) and has been called ‘the soul of mathematics’ (Schoenfeld, 2009), while it has been described as playing “a distinctive role” within the mathematical community, “separat[ing] mathematics from the empirical sciences” (Hoyles, 1997, p. 7). This multifaceted special status of proof within the discipline of mathematics suggests that one might obtain a different picture if one conducted a similar instructionally focused review of other areas in mathematics education research for which content is viewed (for legitimate reasons) more as a context for the research rather than as a core part of it.

Another trend in the recent proof and proving literature that was revealed by our review relates to two C-including categories that appeared to play an important role in the area of research that we considered in this paper: Student-Content (SC) and Student-Teacher-Content (STC). SC was the largest category, which however did not include a high ratio of papers with a breaking ground contribution, while STC had the second highest ratio of such contributing papers. Thus, it may be argued that SC gathers the strongest interest of the researchers who appear to mainly implement and investigate existing ideas and methods (studies more oriented towards building ground), whereas STC along with C seem to be the main areas where new methodological and theoretical perspectives appear (studies more oriented towards breaking ground). An alternative interpretation of this finding is that STC or C are better suited for such theoretical or methodological breaking ground contributions as compared to other categories of the framework. Conducting a similar review of other mathematics education research areas and comparing the findings with ours might cast light on this issue.

Although our review covered a relatively short time period to allow conclusions about evolutionary trends in research on proof and proving, we presume that the STC category, in which about a quarter of the papers in our sample fell, constitutes the focus of a growing body of research in the field, with more researchers attempting to simultaneously address and include instructional aspects in their studies. This rather holistic way of studying instructional practice might reflect a response to Mariotti’s (2006) call in an early review for investigations into the teaching and learning of proof and proving not being divorced from the reality of the classroom, as well as the calls in more recent reviews by Stylianides et al. (2017) and Mariotti et al. (2018) for more research on classroom-based interventions in this area. The publication of an edited volume specifically devoted to classroom-based interventions in the area proof (Stylianides & Stylianides, 2017) might have also helped put this research area in the spotlight. We note further that some themes that may not be dominant in terms of volume of assigned papers appear to transcend the identified categories of instructional focus: notably, the role of technology (within

the learning environment and/or research methodology), linguistic and discursive aspects, proof validation criteria, and affective aspects. We have found that these themes and the holistic, systemic focus (as signified by the STC category) seem to also gain more presence in major conferences in our field and, hence, we posit that they may constitute potential areas of interest for future reviews of research on proof and proving.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11858-023-01518-y>.

Acknowledgements All authors contributed equally to the preparation of this paper. We thank Jisoo Seo who helped with the literature search in her role as research assistant in a proof-related project supported by a John Fell Fund to the first two authors (Grant No: 0005970).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

We mark with **annotated papers and with *papers included in the sample of 103 papers that we considered in the review. A full list of the 103 papers can be found in Supplementary Materials

- **Aaron, W. R., & Herbst, P. G. (2019). The teacher’s perspective on the separation between conjecturing and proving in high school geometry classrooms. *Journal of Mathematics Teacher Education*, 22(3), 231–256. The authors examined secondary teachers’ perspectives on the interplay between conjecturing and proving. Teachers favoured the separation of these two activities, perceiving them as having different goals and drawing on different resources and thus as requiring different actions from students. These teacher perspectives are viewed as an obstacle to students’ engagement in authentic mathematics.
- *Aberdein, A. (2019). Evidence, proofs, and derivations. *ZDM—Mathematics Education*, 51(5), 825–834.
- *Ahmadpour, F., Reid, D., & Reza Fadaee, M. (2019). Students’ ways of understanding a proof. *Mathematical Thinking and Learning*, 21(2), 85–104.
- *Antonini, S. (2019). Intuitive acceptance of proof by contradiction. *ZDM—Mathematics Education*, 51(5), 793–806.
- **Aricha-Metzer, I., & Zaslavsky, O. (2019). The nature of students’ productive and non-productive example-use for proving. *Journal of Mathematical Behavior*, 53, 304–322. The authors conducted

- individual interviews with secondary and undergraduate students on tasks that called for conjecturing and proving. They distinguished between empirical example-use and generic example-use and found a relatively strong tendency among students to use examples generically. The latter was also associated with a productive use of examples in the proving process.
- *Ayalon, M., & Hershkowitz, R. (2018). Mathematics teachers' attention to potential classroom situations of argumentation. *Journal of Mathematical Behavior*, 49, 163–173.
- *Ayalon, M., Wilkie, K. J., & Eid, K. H. (2022). Relating students' emotions during argumentative discourse to their learning of real-life functional situations. *Educational Studies in Mathematics*, 110(1), 23–48.
- *Azrou, N., & Khelladi, A. (2019). Why do students write poor proof texts? A case study on undergraduates' proof writing. *Educational Studies in Mathematics*, 102(2), 257–274.
- Ball, D. L., & Bass, H. (2003). Making mathematics reasonable in school. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A Research companion to principles and standards for school mathematics* (pp. 27–44). National Council of Teachers of Mathematics.
- Baldinger, E. E., & Lai, Y. (2019). Pedagogical context and proof validation: The role of positioning as a teacher or student. *Journal of Mathematical Behavior*, 55, 100698.
- *Bergwall, A. (2021). Proof-related reasoning in upper secondary school: Characteristics of Swedish and Finnish textbooks. *International Journal of Mathematical Education in Science and Technology*, 52(5), 731–751.
- *Blanton, M., Gardiner, A. M., Ristroph, I., Stephens, A., Knuth, E., & Stroud, R. (2022). Progressions in young learners' understandings of parity arguments. *Mathematical Thinking and Learning*, 1–32. <https://doi.org/10.1080/10986065.2022.2053775>
- Boero, P. (2022). Developing students' rationality by constructing and exercising rational mathematical templates: The case of counterexamples. In J. Hodgen, E. Geraniou, G. Bolondi, & F. Ferretti. (Eds.), *Proceedings of the Twelfth Congress of the European Society for Research in Mathematics Education* (pp. 101–108). Free University of Bozen-Bolzano and ERME.
- Boero, P., & Turiano, F. (2023). Proof and students' proving in Euclidean geometry: Linguistic aspects and method issues in the rationality perspective. In A. Schüler-Meyer, J. Ingram, & K. Erath (Eds.), *Proceedings of the Twelfth ERME Topic Conference on Language in the Mathematics Classroom* (pp. 22–29). ERME / HAL Archive.
- *Brown, S. A. (2018). Are indirect proofs less convincing? A study of students' comparative assessments. *Journal of Mathematical Behavior*, 49, 1–23.
- Bruner, J. (1960). *The Process of Education*. Harvard University Press.
- *Buchbinder, O., & McCrone, S. (2020). Preservice teachers learning to teach proof through classroom implementation: Successes and challenges. *Journal of Mathematical Behavior*, 58, 100779.
- *Campbell, T. G., King, S., & Zekowski, J. (2021). Comparing middle grade students' oral and written arguments. *Research in Mathematics Education*, 23(1), 21–38.
- *Cirillo, M., & Hummer, J. (2021). Competencies and behaviors observed when students solve geometry proof problems: An interview study with smartpen technology. *ZDM—Mathematics Education*, 53(4), 861–875.
- Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, 25(2), 119–142.
- **Davies, B., Alcock, L., & Jones, I. (2020). Comparative judgement, proof summaries and proof comprehension. *Educational Studies in Mathematics*, 105(2), 181–197. The authors used comparative judgement to investigate undergraduate students' proof comprehension based on their summaries of a given proof. They showed that comparative judgement can help produce valid and reliable assessments of the quality of students' proof summaries, thereby offering a new way to assess proof comprehension for which there is a scarcity of valid and reliable measures.
- **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. The authors used think-aloud protocols to examine novice and experienced undergraduate students' reading proof processes. They identified three key reading behaviors in which the two groups differed: (1) the types of goals they set for their reading, (2) the way novices did not distinguish propositions in a theorem statement as assumptions or conclusions, and (3) the type and quality of warrants.
- *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), 71–89.
- *Dogan, M. F., & Williams-Pierce, C. (2021). The role of generic examples in teachers' proving activities. *Educational Studies in Mathematics*, 106(1), 133–150.
- Ernest, P. (Ed.). (2018). *The Philosophy of Mathematics Education Today*. Springer International Publishing.
- *Fernández-León, A., Gavilán-Izquierdo, J. M., & Toscano, R. (2021). A case study of the practices of conjecturing and proving of research mathematicians. *International Journal of Mathematical Education in Science and Technology*, 52(5), 767–781.
- *Fujita, T., Jones, K., & Miyazaki, M. (2018). Learners' use of domain-specific computer-based feedback to overcome logical circularity in deductive proving in geometry. *ZDM—Mathematics Education*, 50(4), 699–713.
- **Gabel, M., & Dreyfus, T. (2020). Analyzing proof teaching at the tertiary level using Perelman's new rhetoric. *For the Learning of Mathematics*, 40(2), 15–19. The authors introduced Perelman's new rhetoric to discuss the links between rhetoric and proving in the context of university teaching of proof. They utilized the concept of "the flow of a proof," including the presentation of a proof's logical structure and the role of informal elements in proof presentation, to reveal aspects that are not addressed by other perspectives (notably Toulmin's scheme).
- Goodchild, S., & Sriraman, B. (Eds.). (2012). New perspectives on the didactic triangle: teacher-student-content. *ZDM—Mathematics Education*, 44(5), 581–695.
- *Hanna, G., & Yan, X. (2021). Opening a discussion on teaching proof with automated theorem provers. *For the Learning of Mathematics*, 41(3), 42–46.
- Hanna, G., & Knipping, C. (2020). Proof in mathematics education, 1980–2020: An overview. *Journal of Educational Research in Mathematics*, 30, 1–13.
- Hanna, G., Reid, D., & de Villiers, M. (Eds.). (2019). *Proof Technology in Mathematics Research and Teaching*. Springer.
- Harel, G., & Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. In F. K. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 805–842). Information Age Publishing.
- Hoyle, C. (1997). The curricular shaping of students' approaches to proof. *For the Learning of Mathematics*, 17(1), 7–16.
- *Karpuz, Y., & Atasoy, E. (2020). High school mathematics teachers' content knowledge of the logical structure of proof deriving from figural-concept interaction in geometry. *International Journal of Mathematical Education in Science and Technology*, 51(4), 585–603.
- *Kazemi, E., Ghouseini, H., Cordero-Siy, E., Prough, S., McVicar, E., & Resnick, A. F. (2021). Supporting teacher learning about

- argumentation through adaptive, school-based professional development. *ZDM—Mathematics Education*, 53(2), 435–448.
- *Knuth, E., Zaslavsky, O., & Ellis, A. (2019). The role and use of examples in learning to prove. *Journal of Mathematical Behavior*, 53, 256–262.
- *Ko, Y. Y., & Rose, M. K. (2021). Are self-constructed and student-generated arguments acceptable proofs? Pre-service secondary mathematics teachers' evaluations. *Journal of Mathematical Behavior*, 64, 100912.
- **Komatsu, K., & Jones, K. (2022). Generating mathematical knowledge in the classroom through proof, refutation, and abductive reasoning. *Educational Studies in Mathematics*, 109(3), 567–591. The authors investigated the processes of generation of mathematical knowledge involving refutations in a study with 14–15-year-old students. They paid particular attention to abductive reasoning and *heuristic refutation* (revising conjectures/proofs through discovering and addressing counter-examples). Task design and the role of the teacher emerge as important factors in the discussion.
- *Kontorovich, I. (2021). Minding mathematicians' discourses in investigations of their feedback on students' proofs: A case study. *Educational Studies in Mathematics*, 107(2), 213–234.
- *Lew, K., & Zazkis, D. (2019). Undergraduate mathematics students' at-home exploration of a prove-or-disprove task. *Journal of Mathematical Behavior*, 54, 100674.
- *Lockwood, E., Caughman, J. S., & Weber, K. (2020). An essay on proof, conviction, and explanation: Multiple representation systems in combinatorics. *Educational Studies in Mathematics*, 103(2), 173–189.
- *Lynch, A. G., & Lockwood, E. (2019). A comparison between mathematicians' and students' use of examples for conjecturing and proving. *Journal of Mathematical Behavior*, 53, 323–338.
- Mariotti, M. A. (2006). Proof and proving in mathematics education. In A. Gutiérrez & P. Boero (Eds.), *Handbook of research on the Psychology of Mathematics Education: Past, present and future* (pp. 173–204). Sense Publishers.
- Mariotti, M., Durand-Guerrier, V., & Stylianides, G. J. (2018). Argumentation and proof. In T. Dreyfus, M. Artigue, D. Potari, S. Prediger, & K. Ruthven (Eds.), *Developing Research in Mathematics Education - Twenty Years of Communication, Cooperation and Collaboration in Europe* (pp. 75–89). Routledge.
- **Mejía-Ramos, J. P., & Weber, K. (2020). Using task-based interviews to generate hypotheses about mathematical practice: Mathematics education research on mathematicians' use of examples in proof-related activities. *ZDM—Mathematics Education*, 52(6), 1099–1112. The authors problematized researchers' common use of task-based interviews to gain insights into mathematical practice. They discussed factors that might prevent valid conclusions about authentic mathematical practice based on how mathematicians respond to researcher-generated tasks in laboratory conditions, and they suggested other methods that can complement the use of task-based interviews.
- Moutsios-Rentzos, A. (2022). Proving as multimodal argumentation: an investigation based on Toulmin's scheme. In J. Hodgen, E. Geraniou, G. Bolondi, & F. Ferretti. (Eds.), *Proceedings of the Twelfth Congress of the European Society for Research in Mathematics Education* (pp. 283–290). Free University of Bozen-Bolzano and ERME.
- Moutsios-Rentzos, A., & Plyta, E. (2019). A systemic investigation of students' views about proof in high school geometry: the official and shadow education systems in a school unit. In U. T. Jankvist, M. van den Heuvel-Panhuizen, & M. Veldhuis (Eds.), *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education* (pp. 292–299). Utrecht, the Netherlands: Freudenthal Group & Freudenthal Institute, Utrecht University and ERME.
- Moutsios-Rentzos, A., Kalavasis, F., & Sofos, E. (2017). Learning paths and teaching bridges: The emergent mathematics classroom within the open system of a globalised virtual social network. In G. Aldon, F. Hitt, L. Bazzini, & U. Gellert (Eds.), *Mathematics and Technology* (pp. 371–393). Springer.
- Moutsios-Rentzos, A., & Kalouzoumi-Paizi, F. (2017). Revisiting Odysseus' proving journeys to proof: the 'convergent- bounded' question. In T. Dooley & G. Gueudet (Eds.), *Proceedings of the 10th Congress of European Research in Mathematics Education* (pp. 219–226). Dublin, Ireland: DCU Institute of Education and ERME.
- *Nickel, G. (2019). Aspects of freedom in mathematical proof. *ZDM—Mathematics Education*, 51(5), 845–856.
- *Pinto, A., & Karsenty, R. (2020). Norms of proof in different pedagogical contexts. *For the Learning of Mathematics*, 40(1), 22–27.
- *Pinto, A., & Karsenty, R. (2018). From course design to presentations of proofs: How mathematics professors attend to student independent proof reading. *Journal of Mathematical Behavior*, 49, 129–144.
- *Rogers, K. C., & Kosko, K. W. (2019). How elementary and collegiate instructors envision tasks as supportive of mathematical argumentation: A comparison of instructors' task constructions. *Journal of Mathematical Behavior*, 53, 228–241.
- *Rø, K., & Arnesen, K. K. (2020). The opaque nature of generic examples: The structure of student teachers' arguments in multiplicative reasoning. *Journal of Mathematical Behavior*, 58, 100755.
- Schoenfeld, A. H. (2009). Series editor's foreword: The soul of mathematics. In D. A. Stylianou, M. L. Blanton, & E. J. Knuth (Eds.), *Teaching and Learning Proof Across the Grades: A K-16 Perspective* (pp. xii–xvi). Routledge.
- *Shinno, Y., Miyakawa, T., Iwasaki, H., Kunimune, S., Mizoguchi, T., Ishii, T., & Abe, Y. (2018). Challenges in curriculum development for mathematical proof in secondary school: Cultural dimensions to be considered. *For the Learning of Mathematics*, 38(1), 26–30.
- **Sommerhoff, D., & Ufer, S. (2019). Acceptance criteria for validating mathematical proofs used by school students, university students, and mathematicians in the context of teaching. *ZDM—Mathematics Education*, 51(5), 717–730. The authors employed the same questionnaire to reveal multifaceted convergences and divergences in the acceptance criteria for validating proofs that three conceptually inter-connecting populations hold: school students, undergraduate mathematics students, and mathematicians (ranging from doctoral students to professors).
- *Stewart, S., & Thomas, M. O. (2019). Student perspectives on proof in linear algebra. *ZDM—Mathematics Education*, 51(7), 1069–1082.
- Stylianides, A. J. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education*, 38(3), 289–321.
- Stylianides, A. J. (2016). *Proving in the Elementary Mathematics Classroom*. Oxford University Press.
- Stylianides, A. J., Bieda, K. N., & Morselli, F. (2016). Proof and argumentation in mathematics education research. In A. Gutiérrez, G. C. Leder, & P. Boero (Eds.), *The Second Handbook of Research on the Psychology of Mathematics Education* (pp. 315–351). Sense Publishers.
- Stylianides, A. J., Komatsu, K., Weber, K., & Stylianides, G. J. (2022). Teaching and learning authentic mathematics: The case of proving. In M. Danesi (Ed.), *Handbook of Cognitive Mathematics* (pp. 727–761). Springer Nature.
- **Stylianides, A. J., & Stylianides, G. J. (2022). Introducing students and prospective teachers to the notion of proof in mathematics. *Journal of Mathematical Behavior*, 66, 100957. The authors argued for the importance of designing and implementing a coherent approach to introducing students and prospective teachers to the notion of proof, including seeing a need for proof and developing an operationally functional conceptualization of it. They

- discussed a learning trajectory and a respective instructional sequence to promote it drawing data from two design experiments.
- Stylianides, G. J., & Stylianides, A. J. (Eds.). (2017). Research-based interventions in the area of proof. *Educational Studies in Mathematics*, 96(2), 119–274.
- 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). National Council of Teachers of Mathematics.
- *Tabach, M., Rasmussen, C., Dreyfus, T., & Apkarian, N. (2020). Towards an argumentative grammar for networking: A case of coordinating two approaches. *Educational Studies in Mathematics*, 103(2), 139–155.
- Toerner, G., & Arzarello, F. (2012). Grading mathematics education research journals. *EMS Newsletter*, 86, 52–54.
- *Weber, K., Mejía-Ramos, J. P., & Volpe, T. (2022). The relationship between proof and certainty in mathematical practice. *Journal for Research in Mathematics Education*, 53(1), 65–84.
- *Yee, S. P., Boyle, J. D., Ko, Y. Y. W., & Bleiler-Baxter, S. K. (2018). Effects of constructing, critiquing, and revising arguments within university classrooms. *Journal of Mathematical Behavior*, 49, 145–162.
- *Yopp, D. A., Ely, R., Adams, A. E., Nielsen, A. W., & Corwine, E. C. (2020). Eliminating counterexamples: A case study intervention for improving adolescents' ability to critique direct arguments. *Journal of Mathematical Behavior*, 57, 100751.
- *Zambak, V. S., & Magiera, M. T. (2020). Supporting grades 1–8 pre-service teachers' argumentation skills: Constructing mathematical arguments in situations that facilitate analyzing cases. *International Journal of Mathematical Education in Science and Technology*, 51(8), 1196–1223.
- *Zhuang, Y., & Conner, A. (2022). Secondary mathematics teachers' use of students' incorrect answers in supporting collective argumentation. *Mathematical Thinking and Learning*, 1–24. <https://doi.org/10.1080/10986065.2022.2067932>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.