Research on the teaching and learning of proof: Taking stock and moving forward

Gabriel J. Stylianides, University of Oxford, UK
Andreas J. Stylianides, University of Cambridge, UK
Keith Weber, Rutgers University, USA

Preprint

Introduction

In this chapter we focus on the concept of *proof*, which is important for students' mathematical education but also hard for teachers to teach and difficult for students to learn. In reviewing the research in this area, our thesis in the chapter can be outlined as follows. Research thus far has produced a good empirical basis and a good number of theoretical constructs about how students at different levels understand, typically *mis*understand, proof.[1] Yet, despite progress in this area, proof currently has a marginal place in ordinary mathematics classrooms. More research is needed on how to elevate the role of proof in ordinary classrooms and how to support teachers' work in enhancing students' understandings about proof. We thus argue that more intervention-oriented studies in the area of proof are sorely needed. Intervention-oriented studies should benefit from and capitalize on the substantial theoretical and empirical foundation that is already available.

In the rest of this section, we briefly discuss the importance of proof in mathematics education and we consider the meaning of "proof" in general and in this chapter in particular. We finish with a description of this chapter's scope and organization.

The Importance of Proof in Mathematics Education

The concept of proof has received attention by mathematics education researchers for many decades (e.g., Fawcett, 1938; Kilpatrick, 1922), but more explicitly so in the past few decades. Indeed, there has been an upsurge of publications on various aspects of proof (mathematical, social, cognitive, didactical, philosophical, etc.) in all mathematics education research journals and in books or specialized volumes (e.g., Hanna & de Villiers, 2012; Hanna, Jahnke, & Pulte, 2010; Reid & Knipping, 2010; Stylianou, Blanton, & Knuth, 2010).

While there is some debate on the place and nature of proof within the field of mathematics (e.g., Lakatos, 1976), there is no disagreement that proof is indispensable to the work of mathematicians and to their efforts to deepen mathematical understanding (e.g., Kitcher, 1984). There is also wide recognition that proof should play an important role in all students' mathematical education (e.g., Hanna & Jahnke, 1996; Mariotti, 2006), even in elementary school (e.g., Ball & Bass, 2003; Ball, Hoyles, Jahnke, & Movshovitz-Hadar, 2002; NCTM, 2000; Stylianides, 2007b; Yackel & Hanna, 2003). This is not to suggest that instruction should treat students as "little mathematicians" (Hiebert et al., 1996). Rather, the idea is that, similarly to its role in the field of mathematics, proof is (or should be) indispensable to meaningful engagement with mathematics in school as well (e.g., Hanna, 1990; Hersh, 1993; Mason, 1982).

For example, proof can allow even young children to explore or debate the truth of mathematical assertions based on the logical structure of the mathematical system rather than by appeal to the authority of the teacher or the textbook (e.g., Ball & Bass, 2000, 2003; Lampert, 2001; Reid, 2002; Zack, 1997). Proof can also serve a broad range of functions including the following: verification or falsification (i.e., establishing the truth or falsity of an assertion), explanation (i.e., offering insight into why an assertion is true or false), discovery (i.e., inventing new results), illustrating new methods of deduction, justifying the use of a definition or an axiom system, and communication (i.e., conveying results) (e.g., Bell, 1976; de Villiers, 1990, 1999; Larsen & Zandieh, 2008; Hanna & Barbreau, 2008; Stylianides, 2008a; Weber, 2002, 2010b).

Preprint

It is encouraging to see that curriculum frameworks in different countries are now calling for an important role of proof in the mathematical experiences of all students and as early as the elementary school. Examples can be found in the Common Core State Standards (NGA & CCSSO, 2010) in the United States and in the latest National Mathematics Curriculum in England (Department for Education, 2013). In particular, one of the three core aims that the National Mathematics Curriculum in England sets for students of all ages relates to proof: "[Students should] reason mathematically by following a line of enquiry, conjecturing relationships and generalisations, and developing an argument, justification or proof using mathematical language" (Department for Education, 2013, p. 3).

This statement situates students' engagement with proof in a broader set of activities, such as generalizing and conjecturing, both of which are frequently based on work with specific cases or examples. These activities are critically important as they can support and give meaning to students' engagement with proof (e.g., Bartolini Bussi, 2000; Boero, Garuti, & Mariotti, 1996; Garuti, Boero, & Lemut, 1998; Lannin, Ellis, & Elliott, 2011; Lockwood, Ellis, Dogan, Williams, & Knuth, 2012; Stylianides, 2008a; Zazkis, Liljedahl, & Chernoff, 2008). In this chapter we consider proof within this broader set of activities, which we cluster under the more encompassing activity of *proving*. Yet, we pay particular attention to those aspects of proving that relate to the construction, validation, or comprehension of mathematical arguments intending to meet the standard of proof. Our focus on these aspects was motivated by the fact that they were found to be especially problematic from both teaching and learning points of view.

The Meaning of Proof in Mathematics Education and in This Chapter

Preprint

The term “proof” has been used in a number of different ways in the field of mathematics education. In this section we present some of the definitions of proof discussed in the literature and we finish with the definition of proof we adopt in this chapter.

Some researchers defined proof from a mathematical standpoint, associating it with logical deductions that link premises with conclusions (e.g., Healy & Hoyles, 2001; Knuth, 2002b; Mariotti, 2000a). Some of these definitions also associate proof with one or more of its functions, notably explanation. For example, Knuth (2002b) used proof as “a deductive argument that shows why a statement is true by utilizing other mathematical results and/or insight into the mathematical structure involved in a statement” (p. 86). A related but more formal perspective holds that “a mathematical proof is a formal and logical line of reasoning that begins with a set of axioms and moves through logical steps to a conclusion” (Griffiths, 2000, p. 2).

Other researchers defined proof from cognitive (e.g., Harel & Sowder, 2007) or social (e.g., Balacheff, 1988b; Manin, 1977) standpoints. Harel and Sowder (2007) used the term proof (or justification) to describe “what establishes truth for a person or a community” (p. 806), while reserving the term mathematical proof for “the mathematically institutionalized notion of proof” (p. 807). Balacheff (1988b) defined proof as “an explanation which is accepted by a community at a given time,” where explanation “describe[s] the discourse of an individual intending to establish for somebody else the validity of a statement” (p. 285). This is consistent with Manin’s (1977) view that an argument becomes a proof after the social act of accepting it as a proof.

These different definitions illustrate the lack of consensus about what proof means in mathematics education research. Of course it may not be possible or desirable for all researchers to adopt a common definition of proof, not least because different research goals may be served better by certain definitions than others. Nevertheless, we would benefit as a field from being

Preprint

explicit about the definitions of proof that underpinned our studies (Balacheff, 2002; Reid, 2005). Indeed, that kind of explicitness about our definitions of proof would help interpretation of research findings and facilitate comparisons of findings between different studies, thus contributing also to the development of a more coherent body of research knowledge in this area.

Being consistent with this call for specificity, we state the definition of proof we use herein. In the context of a classroom community at a given time, we understand proof as follows:

> *Proof* is a *mathematical argument*, a connected sequence of assertions for or against a mathematical claim, with the following characteristics:
> 1. It uses statements accepted by the classroom community (*set of accepted statements*) that are true and available without further justification;
> 2. It employs forms of reasoning (*modes of argumentation*) that are valid and known to, or within the conceptual reach of, the classroom community; and
> 3. It is communicated with forms of expression (*modes of argument representation*) that are appropriate and known to, or within the conceptual reach of, the classroom community. (Stylianides, 2007b, p. 291; emphasis in original)

Given this definition of proof, we define *proving* broadly to denote the activity in search for a proof. As we noted earlier, this activity can include a cluster of other related activities that are often precursors to proving, such as generalizing and conjecturing.

While not comprehensive, this definition of proof has certain features that make it appropriate for a review chapter. Specifically, the definition is sufficiently "elastic" to describe proof across all levels of education and, as we explain next, it merges different mathematical, social, cognitive, and didactical points of view. From a mathematical point of view, the definition requires that arguments accepted as proofs use true statements, valid forms of reasoning, and appropriate forms of expression, whereby the terms "true," "valid," and "appropriate" should be understood in the context of contemporary mathematics. From a social point of view, the definition requires that the statements and forms of reasoning or expression that are used in an argument which meets the standard of proof be part of the respective

Preprint

classroom community's shared knowledge or part of knowledge that is potentially accessible to the community. From a cognitive point of view, the definition can be used to describe students' conceptions of proof with respect to each of the three components of an argument identified in the definition: set of accepted statements, modes of argumentation, and modes of argument representation. Finally, from a didactical point of view, the definition can support judgments about whether students' arguments meet the standard of proof and, if not, it can also support decisions about which specific components of students' arguments require development so as to better approximate that standard.

The Chapter's Scope and Organization

While we strove to consider all major strands of research in the area of proof, we doubtlessly and regretfully were unable to discuss all research work that deserved commentary. We acknowledge also that our choices about which research to review have been influenced by a number of factors, including our subjective and inevitably limited understanding of the field, and the desire for the chapter to be concise and coherent.

With regard to the previous point about coherence, our choices of references have been guided by the chapter's specific focus as reflected in the chapter's thesis outlined earlier.[2] As this was a quite different focus from the one in Harel and Sowder's (2007) chapter on proof in the previous Research Handbook, which essentially covered research published up to 2003 and reviewed that research through the lens of the notion of *justification schemes* (to which we return later in the chapter), it was not sensible for us to restrict references in this chapter to post-2003 publications. Furthermore, some foundational work in certain areas of our focus dates before 2003, and so we needed to review some of that older work in order to contextualize more recent research findings and theoretical constructs. Yet, to the extent possible we concentrated on recent

Preprint

publications, notably refereed publications and few doctoral dissertations in important areas with scarcity of published research. Also, we note that a large and important body of research on proof has been published in the refereed proceedings of the International Group for the Psychology of Mathematics Education (PME), but we have not devoted much space in this chapter to discussion of this research as it has been reviewed elsewhere (Mariotti, 2006; Stylianides, Bieda, & Morselli, forthcoming).

The chapter is structured as follows. Next, we review research on proving from three different perspectives that we identified: proving as problem-solving, proving as convincing, and proving as a socially-embedded activity. After that, we explore the place of proof in typical school mathematics classroom practice and discuss factors that contribute to the predominantly marginal place of proof in ordinary classrooms. Then, we review research on classroom-based interventions in the area of proof and discuss possibilities of future interventions in this area.

Research on Proving From Three Different Perspectives

Proving is a multi-faceted activity and educational researchers have investigated it from different perspectives. From our subjective reading of the literature, we identified three broad research perspectives in the area of proof and use those as the organizing structure for our discussion in this section. Specifically, we identified some studies that built upon the problem-solving literature in cognitive psychology and conceptualized proving as *problem-solving*; the aim of research within this perspective is to understand the skills, competencies, and dispositions that students need in order to produce adequate performance on proof-related activities. Alternatively, we identified some studies on proof from a constructivist perspective that conceptualized proving as *convincing*; the aim of research within this perspective is to understand students' or teachers' standards of mathematical conviction and their proximity to

Preprint

acceptable standards in the discipline. Finally, we identified other studies that conceptualized proving as a *socially-embedded activity* and investigated how proof is practiced in mathematical and classroom communities.

In this section we provide an overview of each of these three research perspectives. As the focus of these perspectives has predominantly been on students' or teachers' understandings of proof, we begin our review of each perspective with a summary of desired understandings of proof. We then move on to consider critical questions examined within each perspective, major theoretical constructs related to the perspective, and key empirical findings that relate mostly to students' or teachers' (mis)understandings of proof. We conclude with a critical commentary on the state of research within each perspective and pose questions for future research.

We do not try to give an exhaustive review of all research studies conducted within each perspective. Rather, our focus is on those studies or theoretical constructs that we think have been particularly influential or have the potential to shed substantial light on critical issues in the field. Although we use how each of the three perspectives views proving as an organizing structure for our review, we do not mean to suggest that each study or construct is fully aligned with the respective view of proving in which we discuss it – only that it seems to fit in better with that view as compared to the other views. Likewise, we do not wish to imply that the researchers we cite would generally align themselves with the perspectives in which we classified their particular studies, only that the particular study seemed aligned better with one perspective than others.

Research Within the Proving as Problem-solving Perspective

Desired understandings of proof. Research on proof from a cognitive psychology perspective frequently conceptualizes proving, and proof-related tasks such as evaluating the

Preprint

correctness of a proof, as a special type of problem solving (e.g., Gick, 1986; Koedinger & Anderson, 1990; Koichu, Berman, & Moore, 2007; Mamona-Downs & Downs, 2005; Schoenfeld, 1985; Selden & Selden, 2013; Weber, 2001, 2005). Within the problem-solving perspective, researchers desire that students are able to provide correct, or normatively acceptable, answers to proof-related tasks. Most often, this includes students being able to write arguments that researchers would consider to be proofs. Researchers also investigate whether students can determine if an ostensibly deductive argument is valid or if it contains a logical error (Selden & Selden, 2003). The key focus tends to be on the *process* of producing a proof (and hence the frequent use of the phrases “proof production” and “proof construction”) rather than the meaning of the *product* of this reasoning (i.e., the written proof of the solver). Broader issues of what consists a correct or normatively acceptable answer, or why individuals or communities might wish to engage in these proof-related tasks, do not usually receive emphasis within this perspective but are addressed in the other two perspectives.

Critical questions. Research in the proof as problem-solving perspective frequently involves answering the following questions: (1) What is students’ rate of success on proof-related tasks? (2) What set of competencies (e.g., knowledge, skills, strategies, dispositions) are needed to achieve success on proof-related tasks? (3) How do experts (notably mathematicians) approach proof-related tasks and why are they successful? (4) How do students approach proof-related tasks and what proof-related competencies do they lack? (5) How can we design instruction to help students develop the competencies needed to write and understand proofs successfully? The bulk of the research in this area concerns questions (2), (3), and (4), which are typically investigated by carefully examining the processes of students or mathematicians writing proofs, often using verbal protocol analysis.

Preprint

Theoretical constructs. There are several theoretical constructs that were used to address the aforementioned issues. Selden and Selden (2013) made a key distinction between the *formal-rhetorical* part of proof production and the *problem-centered* part of proof production. Put simply, the formal-rhetorical part of proving involves focusing on the logical aspects of structuring a proof, while the problem-centered aspect of proving involves the creative decision-making aspects of proving. Hence, this dichotomy can be helpful in attributing shortcomings in proof writing to having a faulty understanding of logic or being unable to generate the key insights needed to produce a proof.

The formal rhetorical part of proving involves choosing a *proof framework* (Selden & Selden, 1995) and using the logical structure of the statement to form suitable hypotheses and conclusions and then unpacking the definitions. Selden and Selden (2013) noted that the formal-rhetorical part of proof writing is largely procedural so this might be amenable to direct instruction. As a critical commentary, we wonder if the choice and implementation of a proof framework is entirely procedural and suspect that some aspects of this might fruitfully be understood as an instance of strategic decision-making. For instance, how should a student decide whether to use a direct proof, a proof by contradiction, or a proof by cases?[3] If one decides to use a proof by cases, how would they know what cases to consider (e.g., splitting the integers into odd and even numbers, positive and negative numbers, etc.)? Selden and Selden's (2013) distinction between the formal rhetorical and problem-centered activities in proof writing is valuable, but we suggest the difference between the two is nuanced.

Researchers focusing on the problem-centered part of proving generally aim to understand how the key insights needed to construct a proof are developed. To do so, researchers frequently make a distinction between the processes that an individual uses to produce a proof

Preprint

and the type of formal argument that is the result of these processes. Weber and Alcock (2004, 2009) referred to proofs based entirely on manipulating formal representations of concepts as *syntactic proof productions* and proofs where inferences were drawn from graphs, diagrams, examples, or other informal representations as *semantic proof productions*. Raman (2003) made an analogous distinction between *proofs based on a procedural idea* and *proofs based on a key idea*, with the former being based on manipulation and the latter being based on an insight gained from an informal, private way of understanding a concept.[4] Whereas Weber and Alcock's (2004, 2009) construct is based on an individual's cognition, Raman (2003) argued that a student's propensity to avoid proofs based on key ideas may be due to the epistemological belief that there is no relationship between private ways of knowing and public rigorous proofs.

Garuti, Boero, and Lamut (1998) proposed the related construct of *cognitive unity*, which occurs when there is a continuum between the processes used to generate or evaluate a conjecture and the proof used to verify the conjecture. Whereas "semantic proof productions" and "proofs based on key ideas" involve students forming a bridge between representation systems, "cognitive unity" is centered on bridging the activities of conjecturing and proving. Hence, the aforementioned constructs may be understood as representing similar phenomena (i.e., accounting for the informal reasoning used to produce a proof), but highlight different aspects of the informal reasoning (representation systems, epistemology, and the aim of students' activity).

Taking a critical lens to constructs of this type, we acknowledge that these constructs give much needed attention to the informal aspects of proving. However, their dichotomous nature can cause problems in two respects. First, the line between an informal representation and a formal representation, as in Weber and Alcock's (2009) framework, or whether a particular piece of reasoning is private or public, as in Raman's (2003) framework, seem to be highly

Preprint

dependent on context and may be subjective. Dawkins (2012) argued that verbal descriptions of mathematical phenomena appear to be neither entirely formal nor based on informal representations of concepts, making such statements difficult to situate within Weber and Alcock's semantic/syntactic framework. Second, the dichotomous nature of such frameworks may make them too coarse of a tool to give an accurate account of individual proving episodes (Sandefur, Mason, Stylianides, & Watson, 2013; Weber & Mejia-Ramos, 2009).

Key findings. We will now discuss five key findings of research on the proving as problem-solving perspective.

Students from lower secondary school to the university level often have difficulty writing proofs. This finding relates to the first critical question of this perspective: How successful are students at writing proofs? In a pioneering study notable for its large sample, Senk (1989) asked 1,520 secondary students taking a geometry class to prove four theorems, two of which required only a single deduction beyond the hypotheses. Senk found that only 30% of the students were able to prove at least three theorems and 29% were unable to construct a single proof. More recently, other large-sample studies have found similarly poor success rates with lower secondary school students (e.g., Knuth, Choppin, & Bieda, 2009), high school students (e.g., Healy & Hoyles, 2000), and mathematics majors (e.g., Iannone & Inglis, 2010).

Students often lack many of the competencies needed for proving. The next key findings relate to the competencies needed to write proofs successfully and the extent that students have or lack these competencies. Researchers have posited many reasons that students cannot write proofs. Within Selden and Selden's (2013) formal-rhetorical aspect of proving, students have difficulty choosing a legitimate proof framework to begin their proof (Selden & Selden, 1995), in part because of their difficulty making sense of statements with a complex

Preprint

logical structure, such as statements with nested conditionals (Zandieh, Roh, & Knapp, 2014), or because they do not understand the proof methods they try to use (e.g., Stylianides, Stylianides, & Philippou, 2004, 2007). A particular concern lies with students' misuse of quantifiers. Epp (2009) noted that students have particular difficulties with existence statements, such as not realizing that one should avoid using the same symbol to denote different existential objects. With multiply quantified "for all-there exist" statements, students need to see that the existential variable is dependent on the universal variable. If the universal variable changes, the existential variable will need to change as well (e.g., Arsac & Durrand-Gurrier, 2005).

With regard to the problem-solving aspects of proving, researchers have remarked that students lack proving strategies or heuristics (e.g., Schoenfeld, 1985; Weber, 2001) or are reluctant to use examples or diagrams in their reasoning (Raman, 2003; Weber, 2001). In general, these findings came from small-scale studies that have not been replicated. Confirmatory studies with larger samples, particularly those where the constructs are measured separately from performance on the proving tasks, would place this research on a more secure footing.

The consideration of diagrams and examples is potentially beneficial for students who are writing proofs. In a study of 12 mathematics majors constructing proofs in a real analysis course, Gibson (1998) noted that when these students reached an impasse, they often constructed diagrams. From these diagrams, the students gained insights that helped them overcome their impasse. Gibson noted that "using diagrams helped students complete sub-tasks that they were not able to complete while working with verbal-symbolic representation systems alone" (p. 284) by facilitating understanding, evaluating the truth of statements, generating ideas, and expressing ideas. Sandefur et al. (2013) reported a similar study on example usage where the authors illustrated how small groups of students, working on number theoretic tasks, gained conceptual

Preprint

insights from studying examples that formed the basis for the proofs they wrote. Others have reported case studies showing similar beneficial effects of diagrams and examples (e.g., Alcock & Weber, 2010; Garuti et al., 1998; Lockwood et al., 2012; Pedemonte, 2007; Samkoff, Lai, & Weber, 2012) in proof-writing for both students and mathematicians.

Students often have difficulty translating informal arguments into proofs. Although diagrams and examples sometimes provide students with conceptual insights for why theorems are true, Duval (2007) cautioned that the transition between informal arguments and proofs can be difficult. Just as there are case studies of students successfully translating arguments based on these conceptual insights into proofs, there are also case studies where students are unable to do so (e.g., Alcock & Weber, 2010; Pedemonte, 2007; Pedemonte & Reid, 2011). Alcock and Simpson (2004) found that diagrams can aid proof-writing, but also can inhibit it by providing students with false confidence in conjectures and leaving students confused about what can be assumed and what needs to be proven (echoing concerns by Duval, 2007). In fact, researchers who have compared those who have used examples or diagrams with those who have not, in the learning of advanced mathematics, have found that neither group appears to outperform the other (e.g., Alcock & Inglis, 2008; Alcock & Simpson, 2004, 2005; Pinto & Tall, 1999).

To account for students' difficulty in translating informal arguments into proofs, Pedemonte and her colleagues have conceptualized the *distance* between the informal arguments that students construct and the proofs that could result from those arguments (Pedemonte, 2002, 2007, 2008; Pedemonte & Reid, 2011). For instance, if the *structural distance* is too wide – that is, if the warrants in the informal argument differ considerably from those that would be used in the proof – students will have trouble producing this proof (e.g., Pedemonte, 2007). Zazkis, Weber, and Mejia-Ramos (in press) have identified activities that students who can successfully

Preprint

translate informal arguments into proofs regularly engage in. These include *syntactifying* (expressing informal claims in precise mathematical language), *elaborating* (being explicit about what principles are being used to derive new inferences and justifying facts), and *rewarranting* (providing a logical backing for graphical inferences).

In general, we know that diagrams and examples can be useful for students but often are not. A key aim of future research, then, is understanding what processes can be used to reap the benefits from diagrams and examples and designing instruction so that students can employ diagrams and examples more effectively.

Students and teachers are often unable to distinguish between proofs and invalid arguments. While the majority of research on students' performance on proof-related tasks concerns students' ability to construct proofs, Selden and Selden (2003) argued that students' ability to check a proof for correctness, which they call *validating* a proof, is also important. To measure students' competency at this proof-related task, Selden and Selden (2003) presented eight undergraduates in a transition-to-proof course with four arguments purporting to prove that "if n^2 is divisible by 3, then n is divisible by 3." They found that students initially performed at chance level, only responding correctly about half the time. These findings were replicated in studies with other undergraduates (Alcock & Weber, 2005; Inglis & Alcock, 2012; Ko & Knuth, 2013; Weber, 2010) and were consistent with studies of preservice secondary teachers (Bleiler, Thompson, & Krajcevski, 2014) and inservice secondary teachers (Knuth, 2002a).

There are two competencies that mathematics majors seem to lack that are necessary for proof validation. First, this population will frequently accept a proof of a conditional statement that assumes the conclusion and deduces the antecedent (i.e., proves the converse) (e.g., Inglis & Alcock, 2012; Selden & Selden, 2003; Stylianides et al., 2004; Weber, 2010). That is, these

Preprint

students are frequently unable to identify when an invalid proof framework is used in a proof. Selden and Selden (2003) hypothesized that this was because students focused on local calculations within a proof rather than on its global structure. Second, Weber and Alcock (2005) argued that, to validate a proof, students need to *infer warrants* or the mathematical principles used to deduce new principles. Studies have found that students frequently do not do this (Alcock & Weber, 2005; Inglis & Alcock, 2012) and hence do not recognize when flawed reasoning is being applied.

Critical commentary and directions for future research. While there have been several large scale studies documenting students' difficulties with producing proofs, there is not a widely used instrument that is used to measure students' proof-writing competence. Indeed, there is little overlap between the tasks used in studies that explore students' proof-writing competence. Further, most researchers have tended to focus on what *older students cannot do*, but as we will illustrate in the next section, others emphasize what *younger children can do*. A consequence is that the literature might be interpreted as suggesting that mathematics majors cannot write proofs while young children can. We are skeptical of such a claim; at least it would need empirical support with young children succeeding on a task at which mathematics majors fail. There is also a gap in the literature with older students where we lack a sense of what proof-related tasks they can successfully accomplish. For instance, although we know that mathematics majors struggle to complete non-routine or difficult proofs, we do not know the extent that they could successfully prove more basic theorems, such as "The product of two odd integers is odd" or "If $f(x)$ and $g(x)$ are strictly positive increasing functions, then $f+g(x)$ and $fg(x)$ are also strictly increasing functions." Having a shared measure of proving competence would also allow more

Preprint

insightful comparisons across different studies and different populations, while also providing a means to assess the efficacy of instructional interventions.

Research touting the benefits of using examples or diagrams to facilitate proof writing largely consists of small-scale studies that highlight the affordances of these representations. However, other research failed to find a correlation between students' propensity to use examples and diagrams and their ability to write proofs. Claims that diagrams and examples *have the potential to benefit students* seem fairly well established. What is urgently needed are ways to help students realize the benefits of employing diagrams and examples. Research comparing the ways that successful provers (e.g., mathematicians, strong students) and unsuccessful provers (e.g., typical students) use diagrams and examples is one avenue for addressing this question.

Research into how students *understand* proof is limited (Mejia-Ramos & Inglis, 2009), but many claim students have difficulty understanding the proofs that they read (e.g., Conradie & Firth, 2000; Cowen, 1991). Recently, there have been theoretical advances in what it means to comprehend a proof and how this understanding can be assessed (Mejia-Ramos et al, 2012; Stylianides & Stylianides, 2009a; Yang & Lin, 2008), which may spur more research in this area.

Research Within the Proving as Convincing Perspective

Desired understandings of proof. Within the proving as convincing perspective, the goal of instruction is typically for students to be convinced by the same types of evidence as mathematicians (e.g., Harel & Sowder, 2007). To specify this goal further, one needs to clarify what arguments convince mathematicians. Two assumptions made in much of this literature is that mathematicians are convinced by deductive arguments (i.e., arguments using valid modes of argumentation as specified in our definition of proof) as opposed to arguments based upon empirical evidence and appeals to authority (Harel & Sowder, 1998, 2007; Recio & Godino,

Preprint

2001), and that a proof is an argument that would convince a mathematician or a given mathematical community (Davis & Hersh, 1981; Harel & Sowder, 1998, 2007; Volminik, 1990). Hence, within this perspective, the goal of instruction is for students to recognize the limitations of empirical or authoritative evidence (e.g., Stylianides & Stylianides, 2009b), and to receive high-levels of (or absolute) conviction by deductive arguments.

This perspective differs from the proving as problem-solving perspective in the following sense. Proving as problem-solving frequently focuses on what students can do to produce an acceptable product without necessarily problematizing what an acceptable product is to mathematicians or would be to the students. Proving as convincing tends to focus on students' interpretation of what a personally meaningful and professionally acceptable product would be, while being less concerned with how such a product might be produced.

Critical questions. Research in the proving as convincing perspective often seeks to investigate the types of arguments that convince various groups, such as school or university students and preservice or inservice teachers. Key questions include: (1) What types of arguments or evidence convince various populations of students or teachers that mathematical assertions are true? (2) In what ways do students' or teachers' standards of conviction differ with those of the mathematical community? (3) What types of instruction can encourage students or teachers to better align their ways of obtaining conviction with those held by the mathematical community? The bulk of the research in this area concerns questions (1) and (2), which have typically been investigated using survey or interview methods.

Theoretical constructs. A number of frameworks have been developed to describe, classify, or analyze the types of arguments students or teachers find convincing or the arguments they offer to convince others (e.g., Balacheff, 1988a, 1988b; Bell, 1976; Hadas, Hershkowitz, &

Preprint

Schwarz, 2000; Harel & Sowder, 1998, 2007; Marrades & Gutiérrez, 2000; Miyazaki, 2000; Simon & Blume, 1996). Most of these frameworks were developed in the earlier stages of research on proof and have been used since then by studies on students' or teachers' standards of mathematical conviction. Two frameworks have been particularly influential and have underpinned the development of subsequent frameworks: Balacheff's (1988a, 1988b) hierarchy of arguments and Harel and Sowder's (1998, 2007) framework on justification schemes.

Balacheff's (1998a, 1998b) hierarchy of arguments distinguished between four main argument types, which, as he asserted, "hold a privileged position in the cognitive development of proof" (Balacheff, 1988a, p. 218). These are: (1) *naïve empiricism*, in which a statement is accepted as true based on the confirming evidence offered by examination of a few but not all possible cases; (2) *crucial experiment*, in which a statement is accepted as true based on verification of its truth in one or more carefully selected cases that are considered to put the truth of the statement to the test; (3) *generic example*, in which the truth of a statement is established by means of operations or transformations on a particular case that is considered to be representative of its class; and (4) *thought experiment*, in which the truth of a statement is established by means of internalized action that detaches itself from a particular representation.

Harel and Sowder's (1998, 2007) classification of students' *justification schemes* (or *proof schemes)* signify what arguments convince students and what arguments students offer to convince others. Harel and Sowder organized students' justification schemes into three categories, each of which has several subcategories: (1) *externally-based*, when conviction resides in some external to the student source such as an authority (authoritarian justification scheme), the form of an argument (ritual justification scheme), or meaningless treatment of symbols (non-referential symbolic justification scheme); (2) *empirical*, when conviction is based

Preprint

solely on the use of one or more examples (examples-based or inductive justification scheme) or on perception of one or more drawings (perceptual justification scheme); and (3) *deductive* or *analytic*, when conviction is based on reasoning concerned with the general aspects of a mathematical situation (transformational justification scheme) or on logical deduction of new results from accepted ones (axiomatic justification scheme).

Marrades and Gutiérrez's (2000) framework is an example of a more recent framework that built on and aimed to synthesize earlier frameworks, primarily those by Balacheff, and Harel and Sowder. Also, some other frameworks were specifically designed for use in particular mathematical domains or learning contexts. For example, Miyazaki (2000) noted that the framework he developed would be readily applicable in algebra but not in geometry. Another example of a more specialized framework is by Hadas et al. (2000), which was specifically concerned with student arguments generated in the context of tasks that created contradictory situations and were investigated by students with the mediation of a dynamic geometry tool. One of the argument types was *visual-variations*, which described explanations that use imagery or actual visual variation induced by dragging. According to the researchers, this argument type is particular to explanations in dynamic geometry learning environments.

Taking a critical lens to the available frameworks in this area, we acknowledge that they offer a robust collection of analytic tools for researchers to use or adapt as they study students' or teachers' standards of mathematical conviction. We are not in a position to precisely map certain research goals onto particular frameworks, and indeed it is possible that the same goals might be served reasonably well by different frameworks, as there is clearly overlap among them. We nevertheless observe that the two frameworks we singled out as having been particularly influential for research on students' or teachers' standards of mathematical conviction (Balacheff,

Preprint

1988a, 1988b; Harel & Sowder, 1998, 2007) offer, each one in its own but not necessarily distinctive way, a rather broad coverage of the possible or common arguments that students or teachers can generate or find convincing. The work of Marrades and Gutiérrez (2000) shows further that a synthesis of these two frameworks can be possible thus broadening even more the argument types covered by them. Yet, there may still be situations where researchers can choose to use frameworks with a simpler hierarchical structure (e.g., Bell, 1976; Simon & Blume, 1996) or others with a more specialized focus (e.g., Hadas et al., 2000; Miyazaki, 2000).

While the available frameworks cover a wide range of argument types, we suggest that they are more applicable to arguments generated in the context of proving tasks involving justification of *true* statements over *infinite* sets (e.g., "the sum of any pair of odd numbers is an even number"). These frameworks are less relevant to arguments generated in the context of proving tasks involving refutation of *false* statements (e.g., "all multiples of three are also multiples of six") or statements over *finite* sets (e.g., "there are four prime numbers between 10 and 20") as well as the related proof techniques of disproof by a counterexample and proof by exhaustion. Indeed, existing frameworks would need some expansion or adaptation in order to capture the multitude of arguments (valid and invalid) that might be generated in this broader range of tasks (e.g., Stylianides & Ball, 2008; Tsamir et al., 2008).

Finally, we observe that two constructs discussed in a number of frameworks, albeit with some variation in the used terminology and definitions, have received considerable attention in the relevant literature. The first construct is *empirical argument*, which is generally used to describe an (invalid) argument that purports to show the truth of a statement based on verification of a proper subset of all the cases covered by the statement. According to the definition of proof we use in this chapter (Stylianides, 2007b), empirical arguments fail to meet

Preprint

the standard of proof due to the invalid modes of argumentation they use. Balacheff's (1988a) notions of "naïve empiricism" and "crucial experiment" can be seen as two different kinds of empirical arguments (with the latter being more sophisticated than the former), while Harel and Sowder's (1998) notion of "empirical justification scheme" can be used to describe the conception of a student who considers an empirical argument to be a proof. Empirical arguments received extensive coverage due to students' propensity to produce them.

The second construct is a particularly important type of example-based argument called *generic argument* (also referred to frequently as *generic proof*). A generic argument establishes that a claim about a set of elements holds for a generic element (or example) in that set of elements (Mason & Pimm, 1984). A generic element is an element of that set that possesses no special properties so that the reasoning used for why the claim is true for that element can be generalized to any element. Generic arguments are considered particularly important for two reasons: (1) they can serve as a bridge between empirical arguments and non example-based deductive arguments (e.g., Balacheff, 1988a; Harel, 2001); and (2) due to their concreteness and specificity, generic arguments might be more accessible to students than conventional proofs, especially for students who find the abstract nature of proof to be an intimidating barrier to comprehension (e.g., Leron & Zaslavsky, 2013; Movshovitz-Hadar, 1998; Rowland, 2002).

Key findings. The research studies in this area involved participants from all levels of post-elementary education (including the university and teacher education levels), but they focused mainly on secondary students and preservice elementary teachers. In this section we discuss four key findings from this body of research, which relate to key questions (1) and (2) we listed earlier; findings related to question (3) are discussed later in the chapter.

Preprint

The first two findings concern students' or teachers' understandings of what evidence suffices to justify or refute a statement, with particular attention to the notions of empirical argument, counterexample, and proof. The other two findings concern students' or teachers' understandings of logical principles (notably the equivalence or not between a conditional statement and its converse, inverse, or contrapositive) or specific proof methods (notably proof by mathematical induction). Overall, the research from which these findings derive has identified certain misunderstandings that often lead students or teachers to be convinced by invalid arguments or to be unconvinced by valid arguments and proofs.

Students and teachers are often convinced by empirical arguments as proofs of generalizations. Many studies found that students and teachers were convinced by empirical arguments as proofs of generalizations, as reflected either in their own argument constructions or their evaluations of given arguments (e.g., Buchbinder & Zaslavsky, 2007; Goulding, Rowland, & Barber, 2002; Goulding & Suggate, 2001; Healy & Hoyles, 2000; Knuth, Choppin, Slaughter, & Sutherland, 2002; Morris, 2002, 2007; Sowder & Harel, 2003; Tapan & Arslan, 2009). For example, in an interview study with 34 preservice elementary and lower secondary school teachers in the United States, Morris (2007) found that 41% of the research participants were convinced that an empirical argument of the form of crucial experiment that was presented to them proved, with absolute certainty, that a generalization over an infinite set was true.

Students and teachers are often unconvinced by the power of proof to prove. Some studies found students or teachers to be unconvinced by the power of deductive arguments to establish conclusively the truth of true mathematical generalizations (e.g., Chazan, 1993; Fischbein, 1982; Fischbein & Kedem, 1982; Morris, 2007; Schoenfeld, 1991). For example, in the same study with preservice teachers cited above, Morris (2007) found that only 62% of the

Preprint

participants were convinced that a particular valid argument that met the standard of proof demonstrated, with absolute certainty, that a generalization over an infinite set was true. Some other studies found students or teachers to be unconvinced that a single counterexample can establish conclusively the falsity of a mathematical generalization, and they tended to dismiss the counterexample as a special case (e.g., Balacheff, 1988b; Mason & Klymchuk, 2009; Simon & Blume, 1996; Stylianides et al., 2002). Students' or teachers' limited faith in the power of a single counterexample to refute a generalization is an illustration of the larger set of difficulties that students, including mathematics majors, and teachers tend to have with proof by counterexample (e.g., Ko & Knuth, 2009, 2013; Leung & Lew, 2013).

Students and teachers are often convinced that a conditional statement is equivalent to its converse or inverse, and unconvinced that it is equivalent to its contrapositive. Students and teachers were found to be convinced that a conditional statement is equivalent to its converse (e.g., Hoyles & Küchemann, 2002; Yu, Chin, & Lin, 2004) or to its inverse (e.g., Goetting, 1995; Knuth, 2002a; Stylianides et al., 2004), while they were also found to be unconvinced by a proof of a conditional statement that used the contraposition equivalence rule (e.g., Goetting, 1995; Stylianides et al., 2004). For example, in two large-scale studies with high school students, one in England (Hoyles & Küchemann, 2002) and the other in Taiwan (Yu et al., 2004), more than half of the students considered a conditional statement to be equivalent to its converse, while ten of the participants in a study with 16 American secondary mathematics teachers accepted as a proof an invalid argument that assumed an equivalence between a statement and its inverse (Knuth, 2002a). Also, in a study with 95 Cypriot senior undergraduate students of whom 70 were education majors (preservice elementary teachers) and 25 were mathematics majors, 47% of the education majors and 28% of the mathematics majors rejected an argument that correctly

Preprint

represented a proof by contraposition, while only 76% of the education majors and 60% of the mathematics majors rejected a false conclusion that assumed an equivalence between a conditional statement and its inverse (Stylianides et al., 2004). Students' and teachers' difficulties with conditional statements belong to a broader set of difficulties that include also difficulties related to understanding or using implication (e.g., Durand-Guerrier, 2003; Hoyles & Küchemann, 2002), or formulating and interpreting the negation of a claim as in proof by contradiction (e.g., Antonini, 2001; Antonini & Mariotti, 2008).

Students and teachers are often convinced by superficial features of proof by mathematical induction. University students and teachers were found to have limited understanding of proof by mathematical induction and to be convinced by superficial features of this proof method (e.g., Brown, 2008; Dubinsky, 1986, 1990; Dubinsky & Lewin, 1986; Harel, 2001; Knuth, 2002a; Movshovitz-Hadar, 1993; Smith, 2006; Stylianides et al., 2007). For example, in the same study with American secondary mathematics teachers described earlier (Knuth, 2002a), some of the participants accepted a proof by mathematical induction not because they understood the method, but because they knew the method could be used in similar mathematical situations. Also, in the same study with Cypriot senior undergraduate students described earlier (Stylianides et al., 2007), 47% of the education majors and 35% of the mathematics majors who accepted a proof by mathematical induction of an open sentence were convinced that the truth set of the sentence could not include elements (positive integers) outside of its domain of discourse covered by the proof, i.e., they were not open to the possibility that a proof by mathematical induction might not be as encompassing as it could be.

Critical commentary and directions for future research. A key assumption we made in our review of the literature in this section, and one that has been made (often tacitly) in studies

Preprint

we reviewed, was that the types of arguments students or teachers produce as proofs or evaluate as proofs are indicative of their standards of mathematical conviction. Yet, there is evidence challenging this assumption: it may be easier for solvers to evaluate given arguments than to construct their own arguments (Reiss, Hellmich, & Reiss, 2002); it may be easier for solvers to identify invalid arguments as invalid than to identify valid arguments as valid (Inglis et al., 2013; Reiss et al., 2002); solvers' poor argument constructions can be misleading indicators of what they think would meet the standard of proof, because they may be well aware of the limitations of their arguments but unable to produce better ones (Knuth, Choppin, & Bieda, 2009; Stylianides & Stylianides, 2009a; Weber, 2010; Weber & Mejia-Ramos, in press); students may evaluate given arguments differently based on what would satisfy them personally or what would satisfy their teachers (Healy & Hoyles, 2000). Further, there is the (again often tacit) assumption that students' construction or evaluation of a given empirical or deductive argument is indicative of robust beliefs about empirical or deductive arguments in general. There is also evidence to challenge this assumption: teachers may evaluate given arguments differently based on whether they had previously been exposed to a lesson transcript that included or omitted a student proof (Morris, 2007); and solvers may perform differently in argument construction in the areas of algebra versus geometry (Leung & Lew, 2013) or in argument evaluation when the underpinning logical principle is expressed using words versus symbolic notation (Stylianides et al. 2004). In short, students' and teachers' argument constructions or evaluations may not indicate stable beliefs but appear to depend on context.

The aforementioned findings raise a number of important factors and methodological considerations that future research investigating the types of arguments that convince various groups needs to consider. Also, given that existing studies on the topic considered at most one or

Preprint

two of these factors that can influence inferences about students' or teachers' standards of mathematical conviction, the issue is raised about the different meanings of "conviction" used in or reflected in the findings of those studies (see Weber & Mejia-Ramos, in press). The issue is further complicated by other findings that students' responses to even well-designed survey items can lead, in the absence of follow-up interviews aiming at clarifying students' responses, to conclusions that considerably understate the level of students' understandings of proof (Stylianides & Al-Murani, 2010). Thus, there is a need for the development of more refined methodological approaches to examining students' standards of mathematical conviction and for more consistent application of these approaches in future research on the topic. One recommendation is that participants in these studies should always be given the chance to explain and qualify their responses; by allowing this, one can see if students are absolutely convinced of a claim because of an empirical argument (which can be problematic) or merely think the claim is probably true (which is not problematic) (Weber & Mejia-Ramos, in press). Further, allowing space for explanation can reveal that ostensibly problematic responses by participants may actually be based on rational mathematical reasoning (Stylianides & Al-Murani, 2010).

Finally, while there is extensive research on students' perceptions of empirical arguments, there has been less research on students' perceptions of diagrammatic arguments and mathematicians' standards of conviction. Perhaps one reason mathematicians' justification schemes have received less attention is that it is widely presumed that mathematicians only gain conviction by deductive evidence, rather than also by empirical evidence or appeals to authority. However, recent studies challenge these findings (see Weber, Inglis, & Mejia-Ramos, 2014; Weber & Mejia-Ramos, 2013). Indeed, it is not clear that mathematicians even agree on what

Preprint

constitutes acceptable evidence or a proof (e.g., Dreyfus, 2004; Inglis et al, 2013), which suggests that a more nuanced understanding of mathematical practice is needed.

Research Within the Proving as a Socially-embedded Activity Perspective

The previous perspectives have implicitly treated proving as either an individual activity that is aimed toward solving a problem or as a means of gaining conviction. However, in mathematical practice, proving usually occurs within a social context (i.e., a mathematician engages in proving to justify or explain a claim to their peers and it is the peers who sanction whether an argument is a proof) and it is usually embedded within a broader mathematical activity. Indeed, as we discussed earlier, how and why one proves, and even what constitutes a proof, can be viewed as inextricably linked to the social context in which a proof occurs.

Within this perspective, the emphasis tends to be on *activity* rather than *understandings* (c.f. Sfard, 1998). In particular, a critical point is that individual proof-related tasks (e.g., constructing proofs, reading proofs) are not viewed in isolation (as the other two perspectives often do), but in the context of a broader mathematical activity. If a student or teacher produced a proof, research in this perspective would frequently place emphasis on the meanings of this artifact and how that individual and members of his or her community could subsequently use it.

Desired understandings (or activities). At a broad level, one goal of instruction within this perspective is for students to engage in authentic mathematical activity of proving as it is practiced in the mathematical community, including meaningfully using proof to settle debates about the truth of contentious mathematical assertions (Alibert & Thomas, 1991; Zack, 1997) and as a tool for generating and communicating mathematical knowledge. As Manin noted, good proofs are proofs that make us wiser (cited in Aigner & Schmidt, 1998). A second goal is for classroom communities to use proof for the same reason as mathematicians, including providing

Preprint

explanation (e.g., Hanna, 1990), illustrating new methods to solve problems (e.g., Hanna & Barbreau, 2010), and deepening one's understanding of concepts (e.g., Larsen & Zandieh, 2008).

Critical issues. As this perspective is less developed than the previous two perspectives, research in this area has not coalesced around particular questions. Important issues in this area include: (1) understanding mathematical practice with respect to proof, especially mathematicians' reasons for engaging in proving; (2) identifying what proving is for students and teachers (rather than assuming *a priori* that proving is convincing); (3) designing classroom environments where proof can be seen as a tool for generating and communicating mathematical knowledge; and (4) creating social norms with respect to proof that invite students to prove and provide learning opportunities for students when engaged in proving activities.

Theoretical constructs. Regarding the second research issue above, the most extensive work in this area has been the investigation of high school geometry classrooms by Herbst, Chazan, and their colleagues. Herbst, Nachlieli, and Chazan (2011) observed that most research on teaching has taken the stance that teachers autonomously enact their knowledge and beliefs to attain goals that they personally value. Herbst and Chazan (2003) adopted an alternative perspective, arguing that geometry teachers (and presumably all teachers) have a *practical rationality*, which is a feel for the craft of teaching that consists of: (1) their principles and values that allow them to justify (or discard) possible actions in an instructional situation; and (2) their dispositions (i.e., what values, commitments, and tendencies an outsider might use to describe the teachers' actions). Herbst and Chazan's main findings are that teachers' and students' actions in a geometry classroom are largely shaped by institutional norms specifying what their classroom responsibilities are. (For a review of some of this research, see Herbst et al., 2009).

Preprint

Another strand of research in this area has examined the role of proof within a classroom or mathematical community and how this role might be socially negotiated (e.g., Alibert & Thomas, 1991; Fukawa-Connelly, 2012b). As this body of research is still developing and is not yet coherent, there are not widely used constructs common in this type of research. Yet, available research offers a good basis for further conceptual work in this area.

Key findings. Below we present some promising findings in this area.

Students' perceptions of proof are largely shaped by regularities that students observe in their classrooms. While much of the research on proof has focused on the types of arguments that individual students would find to be personally convincing, students may also evaluate a proof as an artifact that would be judged as acceptable by the mathematical community. Further, Hemmi (2006) hypothesized that, as students may feel that they are only peripheral members of that community, their judgments about the validity of an argument are likely influenced by how similar that argument is to other proofs that they had seen. Hemmi advanced this hypothesis by documenting that undergraduates would infer what counted as a proof by generalizing from the examples of proofs that they had seen in their university classes (see also Drefyus, 1999).

The effects of this student behavior can be seen in studies comparing what arguments students find convincing against what they believe constitutes a proof. For instance, Weber (2010a) presented 28 mathematics majors with a graphical argument in support of the claim that $\int_0^\infty \frac{\sin(x)}{x}\,dx$ was positive. Fourteen participants thought the argument was not a proof, with nine citing the use of a graph as their basis for rejecting this argument. Interestingly, five of the students who thought the argument was not a proof found the argument to be completely convincing, but said the argument could not be a proof because they did not see proofs with graphs in their mathematics classes.

Preprint

The format in which proofs are written can constrain the types of reasoning that take place in mathematics classrooms. At least in the United States, proofs in geometry are sometimes written in a two-column format where a geometry statement appears in the left column and a reason for why that statement is logically permissible appears in the right column (Herbst, 2002a). Herbst (2002b) observed that this format places implicit (and sometimes conflicting) demands on teachers and students that constrain what is possible in a secondary school classroom. For instance, the two-column format implicitly requires giving students a statement (in the form of premises) and a conclusion, and testing students' abilities to reason logically from the statement to the conclusion. This practice discourages students from generating key ideas from a proof, constructing their own diagrams, or making conjectures by choosing their own premises and conclusions (Herbst & Brach, 2006).

Mathematicians usually do not read proofs to gain certainty in theorems but to advance their mathematical agenda. If a theorem or a proof is published in the literature, some mathematicians will not read the proof, but will consider its publication as sufficient grounds to use the theorem in their own research (Auslander, 2008; Geist, Löwe, & Van Kerkhove, 2010; Mejia-Ramos & Weber, 2014; Weber, 2008; Weber & Mejia-Ramos, 2011, 2013). Nonetheless, mathematicians frequently do read proofs, although the most common reason for doing so is to find ideas or techniques that might be applicable in their own work (Weber & Mejia-Ramos, 2011; see also Rav, 1999). This can be seen when one looks at why mathematicians re-prove theorems, which is not usually done to provide conviction, but is often done to provide a more comprehensible explanation for why something is true or to illustrate the power of new techniques or representation systems (Dawson, 2006). The key point here is that, to mathematicians, a published proof is often not the end of an investigation that officially

Preprint

sanctions a theorem; rather the content of the proof is studied by members of the mathematical community to advance their mathematical understanding (Rav, 1999). It is interesting to consider if proof can play an analogous role in mathematics classrooms.

Students and secondary mathematics teachers often do not see proofs as providing explanation and have difficulty understanding proofs. As noted above, a primary purpose of proof for mathematicians is to find new methods for solving problems and advancing one's mathematical understanding. Interviews with mathematicians revealed that they desire that proof play similar roles in the college mathematics courses that they teach (Nardi, 2007; Weber, 2012). However, Knuth's (2002b) interviews with high school teachers showed that they viewed proof as a tool for conviction, not explanation or communication. Large-scale surveys suggest that secondary school students (Healy & Hoyles, 2000) and mathematics majors (Weber & Mejia-Ramos, 2014) appear to view proof in the same limited manner.

Negotiating productive classroom norms can highlight students' responsibilities with respect to proof and thereby create learning opportunities. In analyzing the progression of an undergraduate abstract algebra course, Fukawa-Connelly (2012b) described how the instructor led the class in negotiating classroom norms about students' responsibilities for the presentation of a proof. For instance, the presenters of the proof had the responsibility of explaining and defending their argument, responding to questions and challenges from the audience, explicitly highlighting the high-level ideas of their proof, and only using results that had previously been peer-validated. The audience of the proof had the responsibility of reading the proof carefully, deciding on its correctness, and asking questions when aspects of the proof were unclear. This differed from most traditional university classrooms where, if students presented a proof at all, the presentation was largely formal and the instructor was the one to judge its correctness.

Preprint

Fukawa-Connelly's work illustrated how these norms provided students with opportunities to develop many of the skills that we discussed earlier, under the proving as problem-solving perspective, while also using proof as a means of advancing conceptual understanding. At a more general level, other researchers have illustrated how making it normative that students provide their own justification for a claim – rather than relying on social cues from the teacher (Yackel & Cobb, 1996), an appeal to an authoritative source (Ball & Bass, 2000), or a democratic vote (Weber et al., 2008) – provides students with learning opportunities and autonomy.

Critical commentary and directions for future research. Many researchers have argued that proof should play an explanatory and communicative role in mathematics classrooms (e.g., Knuth, 2002a; Harel & Sowder, 2007). However, as Raman (2003) and Weber (2010b) noted, there is not a consensus amongst mathematics educators as to what it means for a proof to explain; a similar argument can be made about whether a proof fulfills a communicative role. Recent philosophical analyses have revealed that mathematical explanation is a contentious and multifaceted concept (e.g., Mancosu, 2011). Furthermore, empirical studies have found that mathematicians do not agree on informal appraisals of proof on issues such as explanation, simplicity, and clarity (Inglis & Aberdein, in press). If explanation and communication are to be listed as goals of presenting proofs and an analytical tool to examine classroom behavior, mathematics educators will need to operationalize these constructs. Moreover, as mathematicians are heterogeneous with their judgments about these attributes (Inglis & Aberdein, in press), it might not be feasible to fully align these constructs with mathematical practice.

While Herbst and his colleagues have extensively investigated how proof is practiced in high school geometry classrooms and the teachers' rationality for practicing proof in this way (e.g., Herbst et al, 2009), there has been limited research in how proof is practiced in domains

Preprint

such as algebra or in college classrooms. Understanding proving activity in these classrooms and teachers' rationality for supporting these activities would be important topics for future research. Also, it would be important to consider if proof can or should play the same role in classrooms as it does in the mathematical community. As Staples, Bartlo, and Thanheiser (2012) documented, the classroom teacher will sometimes have different professional needs than the mathematician. For instance, a teacher may ask a student to justify a claim that is obvious to a teacher to see if the student really understood the concept in question.

Concluding Remarks

The literature we reviewed in this section generally portrays a gap between desirable and typical student understandings of proof. One might hypothesize that the size of the gap would be inversely proportional to students' age, but the varying methodologies and research instruments used in different studies offer a weak basis on which to examine this hypothesis. Yet, the findings of a large-scale study in England (Küchemann & Hoyles, 2001-03) raised doubt about the possible role of age as the main determining factor of good understanding of proof: the study, which surveyed annually from the end of Year 8 to the end of Year 10 1512 high-attaining students (roughly 13- to 15-year-olds), showed that there were modest (if any) improvements in students' understandings of proof over time. Furthermore, mathematics education research at the elementary school level (e.g., Ball & Bass, 2003; Lampert, 2001; Maher & Martino, 1996; Reid, 2002; Stylianides, 2007a, 2007b, 2007c; Zack, 1997), together with psychological research on the cognitive development of children's ability for deductive reasoning (reviewed in Stylianides & Stylianides, 2008), suggested that some aspects of proof (as defined in Stylianides, 2007b, and used in this chapter) can be conceptually accessible even to young children and that classroom instruction has a crucial role to play in fostering good understanding of proof.

Preprint

Specifically, the aforementioned mathematics education research indicated that, in supportive classroom environments, even young children are capable of writing proofs or developing standards of mathematical conviction that approximate acceptable standards in the mathematical community (see, in particular, Maher & Martino, 1996; Reid, 2002). As most of this research was conducted in classrooms of teacher-researchers or of teachers working closely with researchers, so the research was more about establishing what is *possible* rather than about documenting what *typically* happens in mathematics classrooms. Indeed, as we will discuss next, proof has a marginal place in typical classroom practice.

Proof in Typical School Classroom Practice: Place and Factors That Contribute to it

We begin this section by examining the place of proof in typical K-12 school mathematics classroom practice and conclude that it is marginal.[5] We then discuss different factors that seem to contribute to the marginal place of proof in typical school classroom practice.

The Place of Proof in Typical School Classroom Practice

There is limited research on the place of proof in typical classroom settings. Yet, some useful knowledge about the place of proof in secondary mathematics classrooms internationally comes from the findings of the 1994-1995 International Mathematics and Science Study Video Study (hereafter, TIMSS 1995 Video Study), the more recent 1998-2000 International Mathematics and Science Study Video Study (hereafter, TIMSS 1999 Video Study), and some smaller-scale studies. The TIMSS 1995 Video Study marked the first time when representative samples of teachers from three countries – the United States, Germany, and Japan – were videotaped teaching in secondary mathematics classrooms (Stigler & Hiebert, 1999). The TIMSS 1999 Video Study sampled mathematics lessons in a larger number of countries: Australia, the Czech Republic, Hong Kong SAR, Japan, the Netherlands, Switzerland, and the United States.

Preprint

Manaster (1998) reported an analysis of the content of the 90 lessons in the sample from the three countries that participated in the TIMSS 1995 Video Study. The analysis looked for explicit instances of mathematical reasoning in the lessons, i.e., instances in which mathematical reasoning was clearly stated and argued. Explicit mathematical reasoning was identified in only 22 lessons. All but two of these lessons were in geometry, and all 22 lessons occurred in countries other than the United States. This almost total absence of explicit mathematical reasoning in beginning algebra courses suggested to Manaster that many students may be encouraged to approach algebra in an excessively procedural manner.

Hiebert et al. (2003) analyzed the kinds of mathematical reasoning encouraged by the problems presented in the lessons of the TIMSS 1999 Video Study. A problem was counted as involving "a proof if the teacher or students verified or demonstrated that the result must be true by reasoning from the given conditions to the result using a logically connected sequence of steps" (p. 73). It was found that problems involving proofs were evident to a substantial degree only in Japan: about 39% of the Japanese lessons included at least one proof, and about 26% of the mathematical problems per lesson in Japan included proofs. The corresponding percentages for the other countries were significantly lower.

The marginal place of proof in typical classrooms revealed by the two TIMSS Video Studies was consistent with the findings of two more recent and smaller-scale investigations, which also examined the place of proof in secondary school mathematics classrooms (Bieda, 2010; Sears, 2012; Sears & Chávez, 2014). Bieda's (2010) study included seven middle school mathematics teachers in the United States and is particularly revealing of the place of proof in school mathematics practice, because it examined the treatment of proof in what may be considered to be a best-case scenario: (1) the teacher participants were experienced and well

Preprint

trained users of a textbook series; (2) prior analysis of the textbook series (Stylianides, 2009b) indicated that it was rich in proof-related tasks; and (3) the teachers were observed implementing lessons from the textbook series that were rich in proof-related tasks. About a third of the proof-related tasks in the lessons that Bieda observed were either not implemented in the classroom or were only discussed in small groups. The proof-related tasks that were implemented in the classroom generated a number of opportunities for students to develop generalizations (including conjectures) and offer arguments to justify or refute them. Yet, in about half of these opportunities students did not provide any justification for their generalizations. In the rest of the opportunities the students frequently justified their generalizations with arguments that treated the general case but almost always fell short of meeting the standard of proof, or used arguments (mostly empirical) that did not treat the general case. The teachers in Bieda's study did not seem to provide sufficient feedback to sustain discussions about students' conjectures or justifications, and they were as likely to positively evaluate a non-proof argument as they were a proof.

Sears (2012) and Sears and Chávez's (2014) analysis focused on the place of proof in high school geometry courses, which are the courses traditionally associated with the teaching of proof in the United States. The analysis of the classroom practice of three teachers led to the following conclusions (Sears, 2012, pp. 201-203): (1) teachers were the mathematical authority in the classroom and students were generally encouraged to memorize the teachers' solutions to proof tasks and subsequently follow their example when doing similar proofs; (2) most of the proof tasks related to known facts, which did not encourage students to consider individual cases or make conjectures; and (3) few opportunities were provided for students to understand the importance of doing proofs or make sense of the mathematics involved. These findings are not surprising as they are consistent with how proof has typically been treated in high school

Preprint

geometry courses in the United States (see Stylianides, 2008b, for a historical account).

Factors Contributing to the Marginal Place of Proof in Typical School Classroom Practice

In this subsection, we discuss three possible factors that can account for the marginal place of proof in mathematics classrooms that we described above.

The role of the teacher. Teachers make decisions about what proof-related tasks to implement in their classrooms and how to implement them, thereby playing a crucial role in students' opportunities to learn mathematics in general and proof in particular. The weak knowledge about proof that many teachers have (e.g., Goulding et al., 2002; Harel, 2001; Knuth, 2002a; Morris, 2002, 2007; Movshovitz-Hadar, 1993; Simon & Blume, 1996; Stylianides et al., 2004, 2007), as well as teachers' beliefs, predominantly counterproductive, about proof or the role of proof in students' learning of mathematics (e.g., Bieda, 2010; Furinghetti & Morselli, 2011; Knuth, 2002b; Sears, 2012), influence how teachers treat proof in their classrooms.

The issues related to many teachers' weak knowledge about proof were already discussed earlier. Regarding teachers' beliefs about proof, Knuth (2002b) investigated 17 secondary school mathematics teachers' conceptions of proof from their perspectives as teachers of mathematics. He found that "teachers tended to view proof … as a topic of study rather than as a tool for communicating and studying mathematics" (p. 61) and to consider proof "as an appropriate goal for the mathematics education of a minority of students" (p. 83). These views are consistent with those expressed by teachers in Bieda's (2010) study discussed earlier. Specifically, those teachers tended not to consider proof-related tasks as being important and indicated skepticism of their students' abilities to generate justifications and proofs. They also considered justification "as 'something impressive' or something for students who are developmentally ready" (p. 380).[6] These teachers' beliefs cast some light on their decisions not to implement or discuss a large

Preprint

number of the proof-related tasks in the observed lessons. A similar skepticism about students' ability to productively engage with proof was also indicated in Sears' (2012) study with high school geometry teachers. Teachers believed that proof is a difficult topic for students and explained, on that basis, their custom to encourage students to memorize teachers' solutions to proof tasks and to follow the teachers' example when doing similar proofs.

Yet, even when teachers' knowledge and beliefs are more attuned to teaching proof, there are other major obstacles that teachers face in trying to engage their students in proving. Stylianides, Stylianides, and Shilling-Traina (2013) described how three preservice elementary teachers in the United States tried to engage in proving the students in their mentor teachers' classrooms. Unlike typical preservice elementary teachers, the three had good mathematical knowledge about proof and beliefs that were aligned with reform calls in the United States (e.g., NCTM, 2000) for teachers to make proving central to school mathematics as early as the elementary grades. They developed these knowledge and beliefs as a result of their participation in an intervention study with a focus on proving and problem solving (e.g., Stylianides & Stylianides, 2009b, 2014a, 2014b). Despite this supportive base of relevant knowledge and beliefs, however, the preservice teachers faced the following challenges when implementing proof-related activities in their classrooms: facilitating student work on the tasks without lowering the tasks' cognitive demands, bringing together different student contributions during class discussions, managing time constraints, responding in the moment to students' ideas during a lesson, and managing students' preexisting habits of mind that were not attuned to reasoning mathematically. Cirillo (2011) reported similar findings from a longitudinal interpretive case study of the classroom experiences of a beginning secondary mathematics teacher.

Preprint

The role of curricular resources. Another factor that plays an important role to the place of proof in typical classroom practice concerns the way proof is treated in *curricular resources* (or *curriculum materials*), i.e., materials that teachers use for planning and teaching (for elaboration on these terms see, e.g., Pepin & Gueudet, 2014; Remillard, 2005). Textbooks are the most commonly used kind of curricular resource, being used on average with approximately 75% of fourth- and eighth-grade students internationally (Mullis, Martin, Foy, & Arora, 2012). Indeed, research in this area showed that mathematics textbooks can have a significant influence on the choice and classroom implementation of mathematical tasks, including proving tasks (e.g., Bieda, 2010; Moyer, Cai, Nie & Wang, 2011; Tarr, Chávez, Reys, & Reys, 2006), and thus on students' opportunities to learn (Cai, Ni, & Lester, 2011).

Stylianides (2014, p. 64) discussed how appropriately designed school and teacher education textbooks can offer a leverage point for improving the place of proof in typical classroom practice. Regarding school mathematics textbooks, these could design rich opportunities for students to engage in proving in coherent and consistent ways, while the accompanying teacher guidebooks could offer advice to teachers about mathematical or pedagogical issues related to the classroom implementation of proving tasks. Regarding textbooks used in mathematics teacher education programs, these could help preservice teachers to develop (1) knowledge and beliefs that could productively support their teaching of proof and (2) strategies that could help them deal with the challenges likely to emerge in their teaching of proof (c.f. Cirillo, 2011; Stylianides et al., 2013).

Unfortunately analyses of mathematics textbooks at the elementary school level (Bieda, Ji, Drwencke, & Picard, 2014), the secondary school level (e.g., Davis, Smith, & Roy, 2014; Fujita & Jones, 2014; Otten, Males, & Gilbertson, 2014; Stylianides, 2008c, 2009b; Thompson,

Preprint

Senk, & Johnson, 2012), and the teacher education level (McCrory & Stylianides, 2014) have collectively indicated three disappointing findings. First, there is a limited and inadequately conceptualized place of proof in school mathematics textbooks, in areas such as addressing common student misconceptions and sequencing and distributing proving tasks within and across grade levels. Second, there is inadequate support available in teachers' guidebooks for teachers to implement proving tasks in the classroom, though we do acknowledge that it is an open question how this support should look like (for discussion, see: Cai & Cirillo, 2014; Stylianides, 2008c). Third, there is a limited treatment of proof in textbooks used in mathematics teacher education programs (at least in the United States).

The role of research. The previous discussion paints a bleak picture of the place of proof in typical school mathematics classrooms, especially in United States classrooms where most of the research has focused. This bleak picture is the case even in near optimal conditions, with highly qualified teachers and relatively supportive textbooks. Unfortunately, optimal conditions are uncommon and teachers often lack the knowledge or support to introduce proving tasks effectively. Also, the textbooks typically design few proving opportunities for students and provide little guidance for teachers to support students' engagement with these opportunities.

In addition to factors related to teachers and curricular resources, as discussed earlier, a third factor that contributes to the marginal place of proof in typical classroom practice is the limited research knowledge about how to change the current state of affairs. Specifically, research thus far has offered inadequate support to teachers, teacher educators, and curriculum developers (including textbook authors) about how they might address problems of practice in the area of proof. Also, as Cai and Cirillo (2014, p. 138) pointed out, we lack studies that explore, in connected ways: (1) the opportunities designed for students in mathematics curricular

Preprint

resources to learn about proving; (2) the implementation of those opportunities in the classroom; and (3) students' development of proving skills with the curricular resources used.

Another way in which research could contribute more is by investigating how proof is treated in assessments (school, state, national, etc.). Although there have been studies that examined students' performance on proof-related tasks in various assessments, especially the National Assessment of Educational Progress in the United States (e.g., Arbaugh, Brown, Lynch, & McGraw, 2004; Silver, Alacaci, & Stylianou, 2000), we are unaware of any studies that did a detailed examination of how proof is treated in those assessments. Given the importance that the school systems in many countries assign to students' performance on tests, a more prominent place of proof in assessments could help increase the attention that proof receives in curricular resources and ultimately in the classroom. In relation to assessment, the field is also currently lacking well defined and consistent criteria that could be used both in the development of proving tasks and in evaluating students' performance on those tasks (a separate set of criteria may be needed depending on the particular aspect of proof, or proving, one is interested in).[7]

With the previous comments we do not mean to suggest that there is a total lack of research support about how to improve the place of proof in school mathematics classrooms. Most of the available research that discussed teaching practices rich in students' engagement with proving, however, related to classrooms of teacher-researchers or teachers working in close collaboration with researchers (e.g., Ball & Bass, 2003; Lampert, 2001; Maher & Martino, 1996; Reid, 2002; Stylianides, 2007a, 2007b; Weber, Maher, Powell, & Lee, 2008; Zack, 1997), or involved the use of dynamic geometry environments (e.g., Arzarello, Olivero, Paola, & Robutti, 2002; Baccaglini-Frank & Mariotti, 2010; de Villiers, 2004, 2012; Jones, Gutiérrez, & Mariotti, 2000), which as we explained earlier is the focus of a different chapter in this Handbook. It is

Preprint

still encouraging, though, to see that some researchers have shown how proof can be interwoven into students' activities to advance students' and teachers' mathematical agendas. In the next section we discuss a number of research studies (both at the school and the university levels) that could help improve the place of proof in typical classroom practice.

Classroom-based Interventions in the Area of Proof

We begin our discussion of classroom-based interventions in the area of proof by clarifying our use of key terms, drawing on Stylianides and Stylianides (2013, p. 334). The term *intervention* denotes action taken to improve a situation in relation to the teaching and learning of mathematics, while the term *classroom* is used broadly to denote a formal learning setting at any level of education, from the elementary school to the university including teacher education. Stylianides and Stylianides (2013) observed that, although there have been notable examples of mathematics education research studies on classroom-based interventions since at least the 1930s (Fawcett, 1938), "the number of such studies is small and acutely disproportionate to the number of studies that have documented problems of classroom practice for which solutions are sorely needed" (p. 334). This observation applies to various areas of mathematics education research, and the area of proof is no exception. Yet, as we will discuss in this section, in recent years there has been an upsurge of studies on classroom-based interventions in the area of proof. These intervention studies would not have been possible without the studies that we reviewed in the previous sections, which offered: theoretical constructs and frameworks in the area of proof, a deep understanding of students' and teachers' difficulties with proof, and an appreciation of factors that contribute to the marginal place of proof in typical school classroom practice.

The rest of this section is organized in two parts. In the first we discuss in some detail six studies as examples of recent interventions in the area of proof. Our discussion of each of these

Preprint

studies is more elaborated as compared to our typical discussion of studies in other parts of the chapter, which reflects the importance we assign to this emerging body of interventionist research in the area of proof. In the second part we refer briefly to few other relevant studies that we also consider important but are unable to discuss in detail due to space limitations, and then we discuss directions for future research on classroom-based interventions in the area of proof.

Detailed Discussion of Selected Recent Classroom-based Interventions

The six classroom-based interventions we discuss here were all satisfactorily successful in promoting their intended goals in the area of proof. Those by Mariotti (2000a, 2000b, 2013) and Jahnke and Wambach (2013) focused on secondary school geometry and lasted over extended periods of time (eight lessons or more). The other four focused on the university level. Harel's (2001) intervention took place over a two-week period, whereas the interventions by Stylianides and Stylianides (2009b), Hodds et al. (2004), and Larsen and Zandieh (2008) were all of a much shorter duration (three hours or less).

None of the six intervention studies fits fully within one of the three perspectives of proving we discussed earlier. This was to be expected as an intervention study normally needs to take into account a broad range of factors or goals that cut across perspectives. Yet, it is true that each intervention study comes closer to a particular perspective and less so to the others: Mariotti's and Larsen and Zandieh's studies come closer to proving as a socially-embedded activity, Hodd et al.'s study to proving as problem-solving, and the other three to proving as convincing. We begin with the two interventions that were conducted in secondary classrooms.

Mariotti's (2000a, 2000b) intervention belongs to a broader set of studies about proof in dynamic geometry environments (see, e.g., Jones et al., 2000). The intervention aimed to introduce 15- to16-year-old students in Italy to the deductive approach in geometry. It extended

Preprint

over two years and drew heavily on the fact that the dynamic geometry environment used in the study, Cabri-Géomètre (hereafter, Cabri), enabled the creation of new systems of commands that allowed students to "construct a parallel between the world of Cabri constructions and geometry as a theoretical system" (Mariotti, 2000b, p. 263). The software tools initially available to the students corresponded to the straightedge and compass tools used in the traditional paper-and-pencil environment. As the students developed different geometrical constructions (such as the angle bisector), the Cabri menu was expanded to include new commands (such as the "angle bisector" command) which then became theorems available for use in subsequent constructions. This progressive expansion of the Cabri menu paralleled, according to Mariotti, the enlargement of the theoretical system. Proof served a twofold role: (1) as a tool to ensure the validity of new constructions based on the available commands; and (2) as a key aspect of the social contract established in the classroom according to which constructions had to be justified and accepted by the classroom community before they became theorems. In a more recent publication, Mariotti (2013) elaborated on the crucial role of the teacher in exploiting Cabri's potential to support students overcome key difficulties they faced as they attempted to make the transition from an intuitive to a more deductive approach to geometry. Critical aspects of the teacher's role included purposefully organizing classroom activities and orchestrating whole-class discussions (along the lines discussed in Bartolini Bussi, 1996) around important mathematical content.

Jahnke and Wambach's (2013) intervention with eighth-grade students in Germany focused on the first element of the definition of proof we use in this chapter, namely, the "set of accepted statements," and aimed to foster students' understanding of the fact that the truth of mathematical statements is dependent on the hypotheses and axioms used to derive them. In other words, the intervention aimed to develop students' understanding that a proof (or any other

Preprint

kind of deductive argument) is based upon certain assumptions. The intervention, which extended over eight lessons in the domain of geometry, focused on a historical astronomical situation concerning the attempts of ancient Greeks to model the path of the sun, the so-called "anomaly of the sun." The students were asked to put themselves in the position of the ancient astronomers and to assume that available to them were only the methods and tools that were known at the time. These restrictions in the methods and tools available to the students were similar to the restrictions imposed on the set of software tools available to the students in Mariotti's (2000a, 2000b) study and were found to be an important contributing factor to students becoming more conscious of the role of assumptions in building a deductive theory.

Stylianides and Stylianides (2009b) reported on an intervention that they developed in a four-year design experiment in an undergraduate mathematics course for preservice elementary teachers in the United States. The intervention involved the implementation of a sequence of tasks and lasted less than three hours. Its design relied heavily on two deliberately engineered *cognitive conflicts* (see Zaslavsky, 2005, pp. 299-300, for discussion of the origins of this notion) that motivated stepwise progressions in preservice teachers' knowledge about proof. The progressions in preservice teachers' knowledge culminated in them: (1) recognizing that empirical arguments of any kind, including Balacheff's (1988a, 1988b) naïve empiricism and crucial experiment, offer insecure methods of validating mathematical generalizations; and (2) seeing an *intellectual need* (c.f. Harel, 1998, 2001; Zaslavsky, Nickerson, Stylianides et al., 2012) to learn about secure validation methods (i.e., proofs). The study offered a promising way to begin to address the common misconception that empirical arguments are proofs. In addition, the theoretical framework and the instructional design that underpinned the intervention helped cast some light on how to address a well-documented tendency amongst students: to treat

Preprint

contradictions as exceptions, thereby neither experiencing cognitive conflicts engineered by instruction nor engaging in a process of modifying their understandings so as to resolve the apparent contradictions. Similar to Mariotti's (2000a, 2000b) study, the instructor in Stylianides and Stylianides' (2009b) study (one of the researchers) played a crucial role in facilitating social interactions and in helping preservice teachers resolve the emerging cognitive conflicts and develop their mathematical knowledge (for elaboration on this aspect of the teacher's role, see Stylianides & Stylianides, 2014b). An appropriately adapted version of the intervention was subsequently implemented with similarly promising results in a high attaining secondary mathematics classroom in England taught by the regular classroom teacher (Stylianides, 2009a; Stylianides & Stylianides, 2014b). The successful implementation of the intervention in a new cultural setting, with a different student population, and by a teacher who was not involved in the development of the intervention, offers further support to its design and theoretical underpinning.

Harel (2001) reported on a two-week intervention that was part of a broader teaching experiment he conducted in an elementary number theory course for lower secondary preservice mathematics teachers. The goal of the teaching experiment was to document students' justification schemes and their development during the course. The particular intervention aimed to teach proof by mathematical induction in a way that addressed what Harel described as two major deficiencies of traditional teaching of this proof method. The first deficiency is that the principle of mathematical induction is introduced to students in an abrupt manner that helps them see neither how the principle is born out of a need to solve specific problems nor how it is derived from knowledge the students already have. The second deficiency is that the problems used to introduce and develop students' understanding of the principle of mathematical induction are typically exercises in algebra rather than problems that highlight the need for a recursive

Preprint

argument. Harel's intervention, and his teaching experiment more broadly, were based on a system of pedagogical principles (for details on those principles, see Harel, 2010). According to Harel (2001, p. 206), the most important finding of his intervention study was that "students changed their current ways of thinking, primarily from mere empirical reasoning – in the form of result pattern generalization – into transformational reasoning – in the form of process pattern generalization." The two different kinds of generalization correspond to two distinct ways of thinking about patterns: in the first students focus on regularity in the process of producing a pattern, whereas in the second they focus on regularity in the result.

Hodds et al. (2014) reported on a series of experiments they conducted with undergraduate mathematics students to study the effect of self-explanation training on students' proof comprehension. The training was based on an earlier study by Inglis and Alcock (2012) in which the researchers confirmed using eye-tracking technology two previous findings about undergraduate students' proof comprehension strategies (the first finding was by Selden and Selden, 2003, and the second by Weber, 2010a): In comparison to expert mathematicians, undergraduate students (1) focused considerably more on algebraic features of proofs and less on surrounding text in which logical claims are often made explicit, and (2) read in a more linear manner and made fewer eye movements between the lines of a proof in search of logical relationships. The training in Hodds et al.'s intervention aimed to make progress in addressing the aforementioned limitations in undergraduate students' proof comprehension strategies by focusing students' attention on logical relationships within a proof. Following the positive findings of two experiments under lab conditions, the research team investigated the effect of self-explanation training in a genuine pedagogical setting which showed a similar positive effect that persisted for at least three weeks. Hodds et al. highlighted the following features of their

Preprint

intervention: (1) the self-explanation training is generic and (2) the training takes less than 20 minutes of individual study. These features are indeed key to the implementation of the intervention in the teaching and learning of proof of other undergraduate mathematics students. According to Stylianides and Stylianides (2013, p. 338) such features can help address major obstacles that classroom-based interventions in mathematics education typically find on their way to "scaling up": many of these interventions have long duration and require a substantial commitment from practitioners such as to change their curricula or "reform" their practices.

Larsen and Zandieh's (2008) study is somewhat different from the interventions we have reviewed thus far in the sense that its goal was to make and illustrate a theoretical point that has practical implications. Yet, the study could also be viewed as an intervention and this is the lens through which we discuss it here. Another reason for including this study in our review is that it helps illustrate a set of studies related to proof that explored the utility of the theory of Realistic Mathematics Education (RME) for supporting the learning of undergraduate mathematics (e.g., Rasmussen & King, 2000; Zandieh, Larsen, & Nunley, 2008). Larsen and Zandieh's (2008) study focused on one classroom episode that took place in a university introductory group theory course, taught by Larsen and taken primarily by third year mathematics majors. The design of the course drew heavily on ideas from RME, including Freudenthal's (1991) notion of *guided reinvention*. As Gravemeijer and Doorman (1999, p. 116) explained, the emphasis in guided reinvention "is on the character of the learning process rather than on invention as such," with an important idea being "to allow learners to come to regard the knowledge they acquire as their own private knowledge, knowledge for which they themselves are responsible." The focus of the study was "to explore the utility of recasting Lakatos' (1976) methods of mathematical discovery as heuristics for designing instruction that supports students' reinvention of mathematics"

Preprint

(Larsen & Zandieh, 2008, p. 206). The students in the course worked with definitions and conjectures that were not fully understood by them or were sometimes intentionally imprecisely formulated by the instructor, much like the historical mathematical environments described by Lakatos. As a result, students sometimes simultaneously arrived with a proof and a counterexample to a conjecture. Larsen and Zandieh's study illustrated how students used this apparent paradox to clarify the definition of a concept, refine the conjecture, or analyze the proof for hidden assumptions, which motivated the construction of new mathematical knowledge. In this sense, the production and analysis of arguments (even invalid ones) played an integral role in the creation and modification of concepts, definitions, and conjectures. To conclude, Larsen and Zandieh's (2008) study, together with Weber et al.'s (2008) study we will discuss next, illustrated that in carefully designed learning environments children and university students can engage with proving in ways that can also deepen their understanding of mathematical concepts.

Other Classroom-based Interventions and Directions for Future Research

Due to space limitations we could not discuss more intervention studies to the same level of detail as we discussed the six studies in the previous section. For more comprehensiveness, we refer briefly to other relevant studies that we also consider important, including some older studies that have been influential. We finish with some directions for future research on classroom-based interventions in the area of proof.

Other classroom-based interventions. As we explained earlier, it is unlikely that any intervention study would fit fully within a particular perspective of proving, but it is nevertheless possible to say that a study comes closer to a particular perspective as compared to the other two. Mariotti's (2000a, 2013) and Larsen and Zandieh's (2008) studies were presented earlier as examples of interventions that come closer to the perspective of *proving as a socially-embedded*

Preprint

activity. Some other examples of intervention studies within the same perspective include the following: Perry, Samper, Camargo, Molina, and Echeverry (2009) who carried out a similar intervention as Mariotti in a plane geometry course, but with preservice secondary mathematics teachers; Alibert and Thomas (1991) who described the norms that a calculus class negotiated to settle mathematical arguments aligned with the standards of proof in the mathematical community; and Goos (2004) who illustrated how a teacher established norms and practices that invited high school students to participate in proof.

Hodd et al.'s (2014) study comes closer to the perspective of *proving as problem-solving*. In some other intervention studies within the same perspective researchers attempted to directly teach students processes and heuristics that they believed would lead to successful proof-writing (Anderson, Corbett, Koedinger, & Pelletier, 1995; Weber, 2006; see also Schoenfeld, 1985, who included proof-writing in mathematical problems that he helped students learn how to solve). The study by Anderson et al. (1995) discussed the Geometry Tutor (Anderson, Boyle, & Yost, 1986), a program that helped students master proving processes in geometry classrooms, with students' ability to write proofs improving by more than a standard deviation.

The other three studies we discussed earlier (Harel, 2001; Jahnke & Wambach, 2013; Stylianides & Stylianides, 2009b) come closer to the perspective of *proving as convincing*. Brown's (2014) study with undergraduate mathematics and science students is another example of a study within the same perspective. Brown used some similar tasks to those described in Stylianides and Stylianides (2009b) to investigate potential pathways students can follow in developing skepticism towards empirical evidence as a basis for validation of generalizations, while still being able to use empirical explorations to gain confidence and develop ideas for a proof. It is also worth noting that, in addition to Harel's (2001) study that we reviewed earlier,

Preprint

there are more intervention studies that focused in the particular area of proof by mathematical induction (e.g., Brown, 2008; Ron & Dreyfus, 2004).

Directions for future research. Our discussion in the chapter covered many problems of classroom practice in the area of proof. As a field, we have a good theoretical and empirical research basis to take up the challenge of seeking solutions to more of these problems than we have done thus far. The possibilities for future research on classroom-based interventions in the area of proof are limitless at this point. For example, virtually every issue we discussed in the previous section, about weaknesses in students' or teachers' understandings of proof, calls for the design of an intervention that would aim to address it.

The intervention studies we discussed earlier have taken a step towards addressing some open problems in the teaching and learning of proof. Beyond the specific issues targeted by these interventions, the research knowledge they produced can also be useful in addressing other related issues. For example, the research knowledge acquired from Mariotti (2013) and Stylianides and Stylianides (2009b, 2014b) about the role of the instructor in facilitating class discussions in the area of proof could be instrumental in the design of a new intervention that would aim to help preservice teachers in managing similar classroom discussions. Indeed, successfully orchestrating class discussions is a major challenge not only for preservice teachers who enact proof-related tasks in particular (Stylianides et al., 2013) but also for other teachers who enact challenging tasks more broadly (Stein, Engle, Smith, & Hughes, 2008).

Existing research on classroom-based interventions has barely scratched the surface of few of the open problems in the teaching and learning of proof. Below we offer three examples of issues that we consider particularly important for future research to address.

Preprint

- There are several intervention studies that focused in the particular area of proof by mathematical induction (e.g., Brown, 2008; Harel, 2001; Ron & Dreyfus, 2004). Other proof methods, however, such as proof by contradiction, contraposition, and counterexample have not received similar attention in the literature. Developing students' ability to refute mathematical claims or to show that acceptance of (the negation) of an assertion leads to a contradiction, is important, especially in reform-oriented classrooms where students are encouraged to make conjectures or to engage in discussion and debate the truth of mathematical ideas (e.g., Ball & Bass, 2003; Zack, 1997; Weber et al., 2008). In developing such interventions it would be useful to try and design them in ways that could allow their use to address, concurrently or separately, with some adaptation, student difficulties with related proof methods such as proofs by contraposition and contradiction (see Antonini & Mariotti, 2008). This would not only help achieve efficiency in the teaching of proof methods, but would also help students make connections between those methods and better understand each method's domain of application. A possible way forward in designing such interventions might be to explore whether and how key features of existing interventions on proof by mathematical induction might be applied in the teaching and learning of other proof methods. One such feature might relate to Harel's (1998, 2010) notion of "intellectual need," which was key both to Harel's (2001) intervention on addressing student difficulties related to mathematical induction and Stylianides and Stylianides' (2009b) intervention on addressing the student misconception that empirical arguments are proofs.
- Research at the school level on the teaching of the axiomatic structures of mathematics (which includes research on the notions of axioms, assumptions, definitions, theorems, etc.) has focused almost exclusively on the domain of secondary school geometry and, frequently,

Preprint

in the context of dynamic geometry environments (e.g., Boero et al., 2007; de Villiers, 1998; Jahnke & Wambach, 2013; Mariotti, 2000a, 2000b) though there is also some research at the elementary school level (Stylianides, 2007a). The approaches used by researchers to engage students in the axiomatic structures presented both similarities and differences. A notable difference in some studies relates to whether the starting point of students' engagement should be particular axioms and definitions, or problems that can raise the need for such axioms and definitions. Is one approach preferable to the other? If so, in what kind of instructional settings and in the service of which learning goals? Is it possible to develop a set of principles about engaging students in the axiomatic structures of mathematics that would apply in various mathematical domains (algebra, geometry, etc.) and at various levels of education? Would the same set of principles apply to both technological and non-technological environments?

- There is a great need for interventions to introduce students of different ages to an appropriate set of criteria for judging whether an argument meets the standard of proof. For consistency and coherence in students' learning of proof at different levels of education and in different mathematical domains, this set of proof criteria would have to be sufficiently elastic to be adapted to students' current level of mathematical knowledge but without compromising mathematical accuracy or students' future learning of mathematics. Unless the field designs effective ways to instill in students expandable as well as developmentally and mathematically appropriate views of what a proof means, it is unrealistic to expect that students will make sustained progress in their learning of proof during their mathematical education. The definition of proof we use in this chapter (Stylianides, 2007b), which was originally formulated for use by researchers, has subsequently been adapted and used with

Preprint

secondary students (Stylianides & Al-Murani, 2010, pp. 23-24) and preservice elementary teachers (Stylianides & Stylianides, 2009a, pp. 242-243) but not yet with elementary students. Similar adaptations may also be possible with alternative definitions of proof.

A common theme in the previous discussion had to do with the development of principles or criteria for particular aspects of the teaching of proof (e.g., principles about teaching different proof methods, principles about teaching axiomatic structures, criteria about the meaning of proof) that could be used and tested in various mathematical domains and at various levels of education. In other words, we are setting forth a vision for the development of a unified instructional basis for the teaching of proof. One might argue that searching for such an instructional basis is a chimera, but we think the challenge is worth the effort. Even a small progress in this area would be a major breakthrough and could contribute to the beginning of an ambitious effort for systemic improvement of students' learning of proof.

Conclusion

Research on proof is growing rapidly and this reflects the importance attributed to proof in mathematics education. Taking stock of major achievements of this body of research thus far, we note that there is: (1) a good number of well developed theoretical frameworks casting light on different aspects of the teaching and learning of proof at various levels of education; (2) a rather extended knowledge base about how students understand (typically *mis*understand) proof; (3) a good knowledge base about the current, largely marginal, place of proof in typical school mathematics classroom practice; and (4) a relatively small number of research studies that have developed promising classroom-based interventions to address important issues of the teaching and learning of proof. Accordingly, the bulk of research thus far has focused more on examining,

Preprint

documenting, and understanding the processes underpinning different problems of classroom practice in the area of proof and less on acting upon such problems to generate possible solutions.

We argue that an important step forward would be, then, for the field to capitalize on the robust foundation offered by prior theoretical and empirical work in order to refine existing or design new interventions to alleviate some of the many problems of classroom practice in the area of proof. In order for such interventions to be of most use to the field, they would need to be not only successful in promoting their intended learning outcomes, but also well theorized thus illuminating the mechanisms of success (Stylianides & Stylianides, 2013). Indeed, by clearly articulating the *theoretically essential components* (Yeager & Walton, 2011, p. 288) of an intervention, i.e., the aspects of the intervention presumed to engineer or support the intended learning outcomes, we can enhance the potential of the intervention to be adapted and applied in classrooms other than those that were used for its development. A longer-term goal can be to seek ways to *scale up* (Coburn, 2003) promising interventions, thus coming closer to realizing the vision of giving access to learning opportunities with proof to a larger number of students. Well designed *educative curricular resources*, i.e., resources that aim to promote teacher learning alongside student learning (e.g., Ball & Cohen, 1996; Davis, Palincsar, & Arias, 2014; Davis & Krajcik, 2005; Stylianides, 2008c), have a major role to play in supporting broad dissemination of promising interventions, thus also partly addressing the enduring problem of the large variation in the quality of learning opportunities offered to students in different classrooms (Morris & Hiebert, 2011; Stylianides & Stylianides, 2014b). Of course we do not underestimate the difficulties involved in pursuing such an ambitious research agenda in the area of proof, but we nevertheless believe that the field is now well positioned to take up the challenge.

Acknowledgements

Preprint

We have contributed equally to the preparation of this chapter. We also wish to thank the Editor and four reviewers for their detailed and useful feedback on an earlier version of the chapter.

Notes

[1] Our use of the terms "misunderstandings" or "misconceptions" is intended to designate views that deviate from conventional knowledge. We consider students' (mis)understandings or (mis)conceptions as both starting points and resources for cognitive growth (e.g., Smith, diSessa, & Roschelle, 1993).

[2] We deliberately did not elaborate in this chapter on issues related to proof in dynamic geometry environments, because such issues are discussed in the chapter on geometry in this volume.

[3] Schoenfeld (1980) highlighted using proof by contradiction to prove statements such as "If 2^p-1 is prime, then *p* is prime" as an instance of a shared heuristic used in problem solving.

[4] To avoid misinterpretation, we are not saying that syntactic proof productions or proofs based on a procedural idea are primarily instances of the formal-rhetorical parts of proving. Even if one writes a proof solely by making logical deductions, there can be substantial decision-making, intuition, and creativity in choosing which deductions to make.

[5] Proof plays a greater role in upper-level university mathematics courses, which are typically taught in a "definition-theorem-proof" format (Weber, 2004). In those courses, proof can be considered as the most common pedagogical explanation provided to students (Lai, Weber, & Mejia-Ramos, 2012), with some estimating that about half of a lecture in advanced mathematics is comprised of the instructor presenting proofs (Mills, 2011). An account of mathematics instructors' teaching is beyond the scope of this chapter, but Nardi (2008) gave an extended analysis of a focus group of mathematicians discussing their teaching and a fairly recent review of the relevant literature is found in Fukawa-Connelly (2012a).

[6] It is interesting to note that mathematics professors make similar claims about mathematics majors. In three separate interview studies with mathematics professors on the teaching of proof at the university level, participants indicated that they felt that some mathematics majors simply were not capable of understanding proof (Alcock, 2010; Harel & Sowder, 2009; Weber, 2012). This led some mathematicians in these studies to question whether proof was an appropriate goal for all mathematics majors.

[7] Paolo Boero made a similar point at the "Argumentation and Proof" Thematic Working Group of the 9th Congress of European Research in Mathematics Education (Prague, Czech Republic, 4-8 February, 2015).

Preprint

References

Aigner, M., & Schmidt, V. A. (1998). Interview with Yuri I. Manin: Good proofs are proofs that make us wiser. *The Berlin Intelligencer*, 16-19.

Alcock, L. (2010). Mathematicians' perspectives on the teaching and learning of proof. In F. Hitt, D. Holton, & P. Thompson (Eds.), *Research in Collegiate Mathematics Education VII* (pp. 63-92). Providence, RI: American Mathematical Society.

Alcock, L., & Inglis, M. (2008). Doctoral students' use of examples in evaluating and proving conjectures. *Educational Studies in Mathematics, 69*, 111-129.

Alcock, L. J., & Simpson, A. P. (2004). Convergence of sequences and series: Interactions between visual reasoning and the learner's beliefs about their own role. *Educational Studies in Mathematics, 57*,1-32.

Alcock, L., & Simpson, A. (2005). Convergence of sequences and series 2: Interactions between non-visual reasoning and the learner's beliefs about their own role. *Educational Studies in Mathematics, 58*, 77-110.

Alcock, L., & Weber, K. (2005). Proof validation in real analysis: Inferring and checking warrants. *The Journal of Mathematical Behavior*, *24*(2), 125-134.

Alcock, L., & Weber, K. (2010). Referential and syntactic approaches to proving: Case studies from a transition-to-proof course. *Research in Collegiate Mathematics Education*, *7*, 101-123.

Alibert, D., & Thomas, M. (1991). Research on mathematical proof. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 215-230). Dordrecht, The Netherlands: Kluwer.

Anderson, J. A., Boyle, C., & Yost, G. (1986). The Geometry Tutor. *Journal of Mathematical Behavior*, *17*, 5-20.

Anderson, J. A., Corbett, A., Koedinger, K., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *Journal of the Learning Sciences*, *4*, 167-207.

Antonini, S. (2001). Negation in mathematics: obstacles emerging from an exploratory study. In M. van den Heuvel-Panhuizen (Ed.), *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 49-56). The Netherlands: Utrecht.

Antonini, S., & Mariotti, M. A. (2008). Indirect proof: What is specific to this way of proving? *ZDM – The International Journal on Mathematics Education, 40,* 401-412.

Arbaugh, F., Brown, C., Lynch, K., & McGraw, R. (2004). Students’ ability to construct responses (1992-2000): Findings from short and extended constructed-response items. In P. Kloosterman & F.K. Lester (Eds.), *Results and interpretations of the 1990-2000 Mathematics Assessments of the National Assessment of Educational Progress* (pp. 337- 363). Reston, VA: NCTM.

Arsac, G., & Durand-Guerrier, V. (2005). An epistemological and didactic study of a specific calculus reasoning rule, *Educational Studies in Mathematics*, *60*, 149-172.

Arzarello, F., Olivero, F., Paola, D., & Robutti, O. (2002). A cognitive analysis of dragging practices in cabri environments. *ZDM – The International Journal on Mathematics Education, 34*(3), 66-72.

Auslander, J. (2008). On the roles of proof in mathematics. In B. Gold & R. A. Simons (Eds.) *Proofs and other dilemmas: Mathematics and philosophy* (pp. 61-77). Washington, DC: Mathematical Association of America.

Preprint

Baccaglini-Frank, A., & Mariotti, M.A. (2010). Generating conjectures in dynamic geometry: The maintaining dragging model. *International Journal of Computers for Mathematical Learning, 15*, 225-253.

Balacheff, N. (1988a). Aspects of proof in pupils' practice of school mathematics. In D. Pimm (Ed.), *Mathematics, teachers and children* (pp. 216–235). London: Hodder & Stoughton.

Balacheff, N. (1988b). A study of students' proving processes at the junior high school level. In I. Wirszup & R. Streit (Eds.), *Proceedings of the Second UCSMP International Conference on Mathematics Education* (pp. 284-297). Reston, VA: National Council of Teachers of Mathematics.

Balacheff, N. (2002). The researcher epistemology: a deadlock for educational research on proof. In F. L. Lin (Ed.), *Proceedings of the 2002 International Conference on Mathematics: understanding proving and proving to understand* (pp. 23-44). Taipei, Taiwan: NSC and NTNU. Pre-publication version retrieved November 25, 2011, from www.tpp.umassd.edu/proofcolloquium07/reading/Balachef_Taiwan2002.pdf

Ball, D. L., & Bass, H. (2000). Making believe: The collective construction of public mathematical knowledge in the elementary classroom. In D. Philips (Ed.), *Constructivism in education: Yearbook of the National Society for the Study of Education* (pp. 193-224). Chicago, IL: University of Chicago Press.

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). Reston, VA: National Council of Teachers of Mathematics.

Ball, D. L., & Cohen, D. K. (1996). Reform by the book: What is—or might be—the role of curriculum materials in teacher learning and instructional reform? *Educational Researcher, 25*(9), 6–8.

Ball, D. L., Hoyles, C., Jahnke, H. N., & Movshovitz-Hadar, N. (2002). The teaching of proof. In L. I. Tatsien (Ed.), *Proceedings of the International Congress of Mathematicians* (Vol. III, pp. 907-920). Beijing: Higher Education Press.

Bartolini Bussi, M. G. (1996). Mathematical discussion and perspective drawings in primary school. *Educational Studies in Mathematics, 31,* 11-41.

Bartolini Bussi, M. G. (2000). Early approach to mathematical ideas related to proof making. In P. Boero, P., G. Harel, C. Maher, & M. Miyazaki (Eds.), *Proof and proving in mathematics education, Proceedings of the ICME9 – TSG12*. Japan: Tokyo/Makuhari. Retrieved July 3, 2013, from http://www.lettredelapreuve.it/OldPreuve/ICME9TG12/ICME9TG12Contributions/BartoliniBussiICME00.html

Bell, A. W. (1976). A study of pupil's proof-explanations in mathematical situations. *Educational Studies in Mathematics, 7*(1), 23-40.

Bieda, K. N. (2010). Enacting proof-related tasks in middle school mathematics: challenges and opportunities. *Journal for Research in Mathematics Education, 41*, 351-382.

Bieda, K. N., Ji, X., Drwencke, J., & Picard, A. (2014). Reasoning-and-proving opportunities in elementary mathematics textbooks. *International Journal of Educational Research, 64,* 71-80.

Bleiler, S. K., Thompson, D. R., & Krajčevski, M. (2014). Providing written feedback on students' mathematical arguments: Proof validations of prospective secondary mathematics teachers. *Journal of Mathematics Teacher Education*, *17*, 105-127.

Preprint

Boero, P., Garuti, R., & Lemut, E. (2007). Approaching theorems in grade VIII. In P. Boero (Ed.), *Theorems in school: From history epistemology and cognition to classroom practice* (pp. 249–264). Rotterdam: Sense Publishers.

Boero, P., Garuti, R., & Mariotti, M. A. (1996). Some dynamic mental processes underlying producing and proving conjectures. *Proceedings of the 20th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 121-128). Valencia, Spain.

Brown, S. (2008). Exploring epistemological obstacles to the development of mathematics induction. In M. Zandieh (Ed.) *Proceedings of the 11th Conference for Research on Undergraduate Mathematics Education*, San Diego, CA. (Retrieved on March 26, 2015, from http://sigmaa.maa.org/rume/crume2008/Proceedings/S_Brown_LONG.pdf)

Brown, S. (2014). On skepticism and its role in the development of proof in the classroom. *Educational Studies in Mathematics*, *86*, 311-335.

Buchbinder, O., & Zaslavsky, O. (2007). How to decide? Students' ways of determining the validity of mathematical statements. In D. Pita-Fantasy & G. Philippot (Eds.), *Proceedings of the 5th Congress of the European Society for Research in Mathematics Education* (pp. 561–571), Larnaca: University of Cyprus.

Cai, J., & Cirillo, M. (2014). What do we know about reasoning and proving? Opportunities and missing opportunities from curriculum analyses. *International Journal of Educational Research, 64*, 132-140.

Cai, J., Ni, Y., & Lester, F. K. (2011). Curricular effect on the teaching and learning of mathematics: Findings from two longitudinal studies in China and the United States. *International Journal of Educational Research, 50*, 63–64.

Chazan, D. (1993). High school geometry students' justification for their views of empirical evidence and mathematical proof. *Educational Studies in Mathematics, 24*, 359-387.

Cirillo, M. (2011). "I'm like the Sherpa guide": On learning on teach proof in school mathematics. In B. Ubuz (Ed.), *Proceedings of the 35th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 241-248). Ankara, Turkey: Middle East Technical University.

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher, 32*(6), 3-12.

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

Cowen, C. (1991). Teaching and testing mathematical reading. *American Mathematical Monthly*, *98*, 50-53.

Davis, E. A., & Krajcik, J. S. (2005). Designing educative curriculum materials to promote teacher learning. *Educational Researcher, 34*(3), 3–14.

Davis, E. A., Palincsar, A. S., & Arias, A. M. (2014). Designing educative curriculum materials: A theoretically and empirically driven process. *Harvard Educational Review, 84*(1), 24-52.

Davis, J. D., Smith, D. O., & Roy, A. R. (2014). Reasoning-and-proving in algebra: The case of two reform-oriented U.S. textbooks. *International Journal of Educational Research, 64*, 92-106.

Davis, P., & Hersh, R. (1981). *The mathematical experience*. New York: Viking Penguin Inc.

preprint

Dawkins, P. C. (2012). Extensions of the semantic/syntactic reasoning framework. *For the Learning of Mathematics*, *32*(3), 39-45.

Dawson, J. W. (2006). Why do mathematicians re-prove theorems? *Philosophia Mathematica*, *14*(3), 269-286.

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

de Villiers, M. (1998). To teach definitions in geometry or teach to define? In A. Olivier & K. Newstead (Eds), *Proceedings of the 22nd Annual Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 248-255). Stellenbosch, South Africa.

de Villiers, M. (1999). The role and function of proof. In M. de Villiers (Ed.), *Rethinking proof with the Geometer's Sketchpad* (pp. 3–10). Key Curriculum Press.

de Villiers, M. (2004). Using dynamic geometry to expand mathematics teachers' understanding of proof. *International Journal of Mathematical Education in Science and Technology, 35*, 703-724.

de Villiers, M. (2012). An illustration of the explanatory and discovery functions of proof. *Pythagoras, 33*(3), (8 pages, doi: 10.4102/pythagoras.v33i3.193)

Department for Education. (2013). *Mathematics: Programmes of study: Key Stages 1-2* (National Curriculum in England). Retrieved on October 31, 2013, from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/239129/PRIMARY_national_curriculum_-_Mathematics.pdf.

Dreyfus, T. (1999). Why Johnny can't prove. *Educational Studies in Mathematics*, *38*, 85-109.

Dreyfus, T. (2004). What counts as proof in the mathematics classroom? In M. Kourkoulos, G. Troulis & C. Tzanakis (Eds.), *Proceedings of the 3rd Colloquium on the Didactics of Mathematics* (pp. 114-132). Rethymnon, Crete: Department of Education, University of Crete.

Dubinsky, E. (1986). Teaching mathematical induction I. *Journal of Mathematical Behavior, 5*, 305–317.

Dubinsky, E. (1990). Teaching mathematical induction II. *Journal of Mathematical Behavior, 8*, 285–304.

Dubinsky, E., & Lewin, P. (1986). Reflective abstraction and mathematics education: The generic decomposition of induction and compactness. *Journal of Mathematical Behavior, 5*, 55–92.

Durand-Guerrier, V. (2003). Which notion of implication is the right one? From logical considerations to a didactic perspective. *Educational Studies in Mathematics, 53,* 5-34.

Duval, R. (2007). Cognitive functioning and the understanding of mathematical processes of proof. In P. Boero (Ed.), *Theorems in school: From historic, epistemology, and cognition to classroom practice* (pp. 163-181). Sense Publishers: Rotterdam.

Epp, S. (2009). Proof issues with existential quantification. In F.-L. Lin, F.-J. Hsieh, G. Hanna, & M. de Villiers (Eds.), *Proceedings of ICMI Study 19: Proof and proving in mathematics education* (Vol. 1, pp. 154-159). Taipei, Taiwan.

Fawcett, H. P. (1938). *The nature of proof (1938 Yearbook of the National Council of Teachers of Mathematics).* New York, NY: Bureau of Publications, Teachers College, Columbia University.

Fischbein, E. (1982). Intuition and proof. *For the Learning of Mathematics*, *3*(2), 9-18.

Preprint

Fischbein, E., & Kedem, I. (1982). Proof and certitude in the development of mathematical thinking. In A. Vermandel (Ed.), *Proceedings of the 6th International Conference of the Psychology of Mathematics Education* (pp. 128–131). Antwerp, Belgium.

Freudenthal, H. (1991). *Revisiting mathematics education.* Dordrecht: Kluwer Academic Publishers.

Fujita, T., & Jones, K. (2014). Reasoning-and-proving in geometry in school mathematics textbooks in Japan. *International Journal of Educational Research, 64*, 81-91.

Fukawa-Connelly, T. (2012a). A case study of one instructor's lecture-based teaching of proof in abstract algebra: making sense of her pedagogical moves. *Educational Studies in Mathematics*, *81*(3), 325-345.

Fukawa-Connelly, T. (2012b). Classroom sociomathematical norms for proof presentation in undergraduate in abstract algebra. *Journal of Mathematical Behavior*, *31*(3), 401-416.

Furinghetti, F., & Morselli, F. (2011). Beliefs and beyond: Hows and whys in the teaching of proof. *ZDM – The International Journal on Mathematics Education, 43,* 587-599.

Garuti, R., Boero, P., & Lemut, E. (1998). Cognitive unity of theorems and difficulty of proof. In A. Olivier & K. Newstead (Eds.), *Proceedings of the 22nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 345-352). Stellenbosch, South Africa.

Geist, C., Löwe, B., & Van Kerkhove, B. (2010). Peer review and testimony in mathematics. In B. Löwe & T. Müller (Eds.), *Philosophy of mathematics: Sociological aspects and mathematical practice* (pp. 155-178). London: College Publications.

Gibson, D. (1998). Students' use of diagrams to develop proofs in an introductory real analysis. *Research in Collegiate Mathematics Education, 2*, 284-307.

Gick, M. (1986). Problem-solving strategies. *Educational Psychologist*, *21*, 99-120.

Goetting, M. (1995). *The college students' understanding of mathematical proof.* Unpublished doctoral dissertation, University of Maryland, College Park.

Goos, M. (2004). Learning mathematics in a classroom community of inquiry. *Journal for Research in Mathematics Education*, *35*, 258-291.

Goulding, M., & Suggate, J. (2001). Opening a can of worms: Investigating primary teachers' subject knowledge in mathematics. *Mathematics Education Review, 13,* 41-44.

Goulding, M., Rowland, T., & Barber, P. (2002). Does it matter? Primary teacher trainees' subject knowledge in mathematics. *British Educational Research Journal, 28*, 689–704.

Gravemeijer, K., & Doorman, M. (1999). Context problems in realistic mathematics education: A calculus course as an example. *Educational Studies in Mathematics, 39*, 111-129.

Griffiths, P. A. (2000). Mathematics at the turn of the millennium. *American Mathematical Monthly, 107,* 1-14.

Hadas, N., Hershkowitz, R., & Schwarz, B. B. (2000). The role of contradiction and uncertainty in promoting the need to prove in dynamic geometry environments. *Educational Studies in Mathematics, 44,* 127-150.

Hanna, G. (1990). Some pedagogical aspects of proof. *Interchange, 21*(1), 6-13.

Hanna, G. & Barbeau, E. (2008). Proofs as bearers of mathematical knowledge. *ZDM – The International Journal on Mathematics Education, 40*, 345-353.

Hanna, G., & Barbeau, E. (2010). Proofs as bearers of mathematical knowledge. In G. Hanna, H .N. Jahnke, & H. Pulte (Eds.), *Explanation and proof in mathematics* (pp. 85-100). Dordrecht: Springer.

Preprint

Hanna, G., & de Villiers, M. (Eds.). (2012). *Proof and proving in mathematics education: The 19th ICMI Study.* Springer.

Hanna, G., & Jahnke, H. N. (1996). Proof and proving. In A. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 877-908). Dordrecht: Kluwer Academic Publishers.

Hanna, G., Jahnke, H. N., & Pulte, H. (Eds.). (2010). *Explanation and proof in mathematics: Philosophical and educational perspectives.* Springer.

Harel, G. (1998). Two dual assertions: The first on learning and the second on teaching (or vice versa). *The American Mathematical Monthly, 105*, 497–507.

Harel, G. (2001). The development of mathematical induction as a proof scheme: A model for DNR-based instruction. In S. Campbell & R. Zaskis (Eds.), *Learning and teaching number theory: Research in cognition and instruction* (pp. 185-212). Dordrecht, The Netherlands: Kluwer.

Harel, G. (2010). DNR-based instruction in mathematics as a conceptual framework. In S. Barath & L. English (Eds.), *Theories of mathematics education* (pp. 343-367). Springer.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. H. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Research in collegiate mathematics education III* (pp. 234–283). Providence, RI: American Mathematical Society.

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). Greenwich, CT: Information Age Publishing.

Harel, G., & Sowder, L. (2009). College instructors' views of students vis-à-vis proof. In M. Blanton, D. Stylianou, & E. Knuth (Eds.), *Teaching proof across the grades: A K-12 perspective* (pp. 275-289). New York, NY: Routledge.

Healy, L., & Hoyles, C. (2000). Proof conceptions in algebra. *Journal for Research in Mathematics Education, 31*, 396-428.

Healy, L., & Hoyles, C. (2001). Software tools for geometrical problem solving: Potentials and pitfalls. *International Journal of Computers for Mathematical Learning, 6*, 235-256.

Hemmi, K. (2006). *Approaching proof in a community of mathematical practice*. Doctoral dissertation. Stockholm University. (Retrieved on June 6, 2015, from http://www.diva-portal.org/smash/get/diva2:189608/FULLTEXT01.pdf)

Herbst, P. G. (2002a). Establishing a custom of proving in American school geometry: Evolution of the two-column proof in the early twentieth century. *Educational Studies in Mathematics*, *49*, 283-312.

Herbst, P. G. (2002b). Engaging students in proving: A double bind on the teacher. *Journal for Research in Mathematics Education*, *33*(3), 176-203.

Herbst, P., & Brach, C. (2006). Proving and doing proofs in high school geometry classes: What is it that is going on for students? *Cognition and Instruction*, *24*(1), 73-122.

Herbst, P., & Chazan, D. (2003). Exploring the Practical Rationality of Mathematics Teaching through Conversations about. *For the learning of Mathematics*, *23*(1), 2-14.

Herbst, P., Nachlieli, T., & Chazan, D. (2011). Studying the practical rationality of mathematics teaching: What goes into "installing" a theorem in geometry? *Cognition and Instruction*, *29*(2), 218-255.

Herbst, P. Chen, C., Weiss, M., & González, G. (with Nachlieli, T., Hamlin, M., and Brach, C.) (2009). "Doing proofs" in geometry classrooms. In M. Blanton, D. Stylianou, & E.

Knuth (Eds.), *Teaching and learning of proof across the grades: A K-16 perspective* (pp. 250-268). New York: Routledge.

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

Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K., Human, P., Murray, H., Olivier, A., & Wearne, D. (1996). Problem solving as a basis for reform in curriculum and instruction: The case of mathematics. *Educational Researcher, 25*(4), 12–21.

Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., Jacobs, J., Chui, A. M., Wearne, D., Smith, M., Kersting, N., Manaster, A., Tseng, E., Etterbeek, W., Manaster, C., Gonzales, P., & Stigler, J. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 Video Study.* NCES 2003-013. Washington, DC: U.S. Department of Education, Institute of Education Sciences.

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

Hoyles, C., & Küchemann, D. (2002). Students' understanding of logical implication. *Educational Studies in Mathematics, 51*, 193–223.

Iannone, P., & Inglis, M. (2010). Self-efficacy and mathematical proof: Are undergraduates good at assessing their own proof production ability? In, *Proceedings of the 13th Conference for Research in Undergraduate Mathematics Education*. Raleigh, North Carolina.

Inglis, M., & Aberdein, A. (in press). Diversity in proof appraisal. To appear in B. Larvor (Ed.), *Mathematical Cultures*. Birkhäuser Science.

Inglis, M., & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education*, *43*, 358-390.

Inglis, M., Mejia-Ramos, J. P., Weber, K., & Alcock, L. (2013). On mathematicians' different standards when evaluating elementary proofs. *Topics in Cognitive Science*, *5*, 270-282.

Jahnke, H. N., & Wambach, R. (2013). Understanding what a proof is: a classroom-based approach. *ZDM – The International Journal on Mathematics Education*, *45*, 469-482.

Jones, K., Gutiérrez, A., & Mariotti, M. A. (Eds.). (2000). Proof in dynamic geometry environments. *Educational Studies in Mathematics, 44*(1/2).

Kilpatrick, W. H. (1922). The next step in method. *The Mathematics Teacher, 15*(1), 16-25.

Kitcher, P. (1984). *The nature of mathematical knowledge.* New York, NY: Oxford University Press.

Knuth, E. J. (2002a). Secondary school mathematics teachers' conceptions of proof. *Journal for Research in Mathematics Education, 33*, 379-405.

Knuth, E. J. (2002b). Teachers' conceptions of proof in the context of secondary school mathematics. *Journal of Mathematics Teacher Education, 5*(1), 61-88.

Knuth, E. J., Choppin, J., & Bieda, K. (2009). Middle school students' productions of mathematical justification. In M. Blanton, D. Stylianou, & E. Knuth (Eds.) *Teaching and Learning Proof Across the Grades: A K-16 Perspective* (pp. 153-170). Routledge: New York.

Knuth, E. J., Choppin, J., Slaughter, M., & Sutherland, J. (2002). Mapping the conceptual terrain of middle school students' competencies in justifying and proving. In D. S. Mewborn, P. Sztajn, D. Y. White, H. G. Weigel, R. L. Bryant, & K. Nooney (Eds.), *Proceedings of the 24th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 1693-1670). Athens, GA: Clearinghouse for Science, Mathematics, and Environmental Education.

Preprint

Ko, Y. Y., & Knuth, E. J. (2009). Undergraduate mathematics majors' writing performance producing proofs and counterexamples about continuous functions. *The Journal of Mathematical Behavior, 28*(1), 68-77.

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

Koedinger, K., & Anderson, J. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science*, *14*, 511-550.

Koichu, B., Berman, A., & Moore, M. (2007). Heuristic literacy development and its relation to mathematical achievements of middle school students. *Instructional Science*, *35*(2), 99-139.

Küchemann, D., & Hoyles, C. (2001–03). *Longitudinal Proof Project* (Technical reports for Year 8–10 surveys). London: Institute of Education. Retrieved on May 3, 2009, from http://www.mathsmed.co.uk/ioe-proof/techreps.html.

Lai, Y., Weber, K., & Mejía-Ramos, J. P. (2012). Mathematicians' perspectives on features of a good pedagogical proof. *Cognition and Instruction*, *30*(2), 146-169.

Lakatos, I. (1976). *Proofs and refutations: The logic of mathematical discovery.* Cambridge University Press.

Lampert, M. (2001). *Teaching problems and the problems of teaching*. Yale University Press.

Lannin, J., Ellis, A. B., & Elliott, R. (2011). *Essential understandings project: Mathematical reasoning (Gr. K-8).* Reston, VA: National Council of Teachers of Mathematics.

Larsen, S., & Zandieh, M. (2008). Proofs and refutations in the undergraduate mathematics classroom. *Educational Studies in Mathematics*, *67*(3), 205-216.

Leron, U., & Zaslavsky, O. (2013). Generic proving: reflections on scope and method. *For the Learning of Mathematics, 33*(3), 24-30.

Leung, I. K. C., & Lew H. (2013). The ability of students and teachers to use counter-examples to justify mathematical propositions: a pilot study in South Korea and Hong Kong. *ZDM – The International Journal on Mathematics Education, 45,* 91-105.

Lockwood, E., Ellis, A. B., Dogan, M.F., Williams, C., & Knuth, E. (2012). A framework for mathematicians' example-related activity when exploring and proving mathematical conjectures. In L. R. Van Zoest, J. J. Lo, & J. L. Kratky (Eds.), *Proceedings of the 34th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 151–158). Kalamazoo, MI: Western Michigan University.

Maher, C. A., & Martino, A. M. (1996). The development of the idea of mathematical proof: A 5-year case study. *Journal for Research in Mathematics Education, 27,* 194-214.

Mamona-Downs, J., & Downs, M. (2005). The identity of problem solving. *The Journal of Mathematical Behavior*, *24*(3), 385-401.

Manaster, A. B. (1998). Some characteristics of eighth grade mathematics classes in the TIMSS videotape study. *American Mathematical Monthly, 105*, 793-805.

Mancosu, P. (2011). Explanation in mathematics. In *Stanford Encyclopedia of Philosophy (on-line)*. (Retrieved on May 19, 2015, from http://seop.illc.uva.nl/entries/mathematics-explanation/.)

Manin, Y. I. (1977). *A course in mathematical logic.* New York, NY: Springer-Verlag.

Mariotti, M. A. (2000a). Introduction to proof: The mediation of a dynamic software environment. *Educational Studies in Mathematics*, *44*, 25-53.

Draft

Mariotti, M. A. (2000b). Justifying and proving in the Cabri environment. *International Journal of Computers for Mathematical Learning, 6*, 257-281.

Mariotti, M. A. (2006). Proof and proving in mathematics education. In A. Gutiérrez & P. Boero (Eds.), *Handbook of research on the PME: Past, present and future* (pp. 173-204). Rotterdam, The Netherlands: Sense Publishers.

Mariotti, M. A. (2013). Introducing students to geometric theorems: how the teacher can exploit the semiotic potential of a DGS. *ZDM – The International Journal on Mathematics Education*, *45*, 441-452.

Marrades, R., & Gutiérrez, Á. (2000). Proofs produced by secondary school students learning geometry in a dynamic computer environment. *Educational Studies in Mathematics, 44*, 87-125.

Mason, J., (with Burton, L., & Stacey, K.). (1982). *Thinking mathematically.* London: Addison-Wesley.

Mason, J., & Klymchuk, S. (2009). *Using counter-examples in calculus.* London: Imperial College Press.

Mason, J., & Pimm, D. (1984). Generic examples: Seeing the general in the particular. *Educational Studies in Mathematics, 15*, 277–289.

McCrory, R., & Stylianides, A. J. (2014). Reasoning-and-proving in mathematics textbooks for prospective elementary teachers. *International Journal of Educational Research, 64,* 119-131.

Mejia-Ramos, J. P., & Inglis, M. (2009). Argumentative and proving activities in mathematics education research. In F.-L. Lin, F.-J. Hsieh, G. Hanna, & M. de Villiers (Eds.), *Proceedings of the ICMI study 19 conference: Proof and proving in mathematics education* (Vol. 2, pp. 88-93). Taipei, Taiwan.

Mejia-Ramos, J. P., & Weber, K. (2014). Why and how mathematicians read proofs: further evidence from a survey study. *Educational Studies in Mathematics*, *85*(2), 161-173.

Mejia-Ramos, J. P., Fuller, E., Weber, K., Rhoads, K., & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics*, *79*(1), 3-18.

Mills, M. (2011). Mathematicians' pedagogical thoughts and practices in proof presentation. Presentation at the *14th Conference for Research in Undergraduate Mathematics Education.* (Abstract retrieved on June 6, 2015, from http://sigmaa.maa.org/rume/crume2011/Preliminary_Reports.html.)

Miyazaki, M. (2000). Levels of proof in lower school mathematics. *Educational Studies in Mathematics, 41*, 47-68.

Morris, A. K. (2002). Mathematical reasoning: Adults' ability to make the inductive-deductive distinction. *Cognition and Instruction, 20*(1), 79–118.

Morris, A. K. (2007). Factors affecting pre-service teachers' evaluations of the validity of students' mathematical arguments in classroom contexts. *Cognition and Instruction, 25*(4), 479-522.

Morris, A. K., & Hiebert, J. (2011). Creating shared instructional products: An alternative approach to improving teaching. *Educational Researcher, 40*(5), 5–14.

Movshovitz-Hadar, N. (1988). Stimulating presentation of theorems followed by responsive proofs. *For the Learning of Mathematics, 8*(2), 12–19, 30.

Movshovitz-Hadar, N. (1993). The false coin problem, mathematical induction and knowledge fragility. *Journal of Mathematical Behavior, 12*(3), 253-268.

Preprint

Moyer, J. C., Cai, J., Nie, B., & Wang, N. (2011). Impact of curriculum reform: Evidence of change in classroom instruction in the United States. *International Journal of Educational Research, 50*(2), 87-99.

Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics.* Boston: Boston College/IEA.

Nardi, E. (2008). *Amongst mathematicians: Teaching and learning mathematics at university level*. USA: Springer.

National Council of Teachers of Mathematics (NCTM) (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

National Governors Association Center for Best Practices, & Council of Chief State School Officers (NGA & CCSSO) (2010). *Common core state standards mathematics*. Washington, DC: Authors.

Otten, S., Males, L. M., & Gilbertson, N. J. (2014). The introduction of proof in secondary geometry textbooks. *International Journal of Educational Research, 64,* 107-118.

Pedemonte, B. (2002). Relation between argumentation and proof in mathematics: cognitive unity or break. In J. Novotná (Ed.) *Proceedings of the 2nd Conference of the European Society for Research in Mathematics Education* (Vol. 2, pp. 70-80). Czech Republic.

Pedemonte, B. (2007). How can the relationship between argumentation and proof be analysed? *Educational Studies in Mathematics*, *66*, 23-41.

Pedemonte, B. (2008). Argumentation and algebraic proof. *ZDM – The International Journal on Mathematics Education*, *40*, 385-400.

Pedemonte, B., & Reid, D. (2011). The role of abduction in proving processes. *Educational Studies in Mathematics*, *76*, 281-303.

Pepin, B., & Gueudet, G. (2014). Curriculum resources and textbooks in mathematics education. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 132-135). Berlin, Heidelberg: Springer.

Perry, P., Samper, C., Camargo, L., Molina, Ó., & Echeverry, A. (2009). Assigning mathematics tasks versus providing pre-fabricated mathematics in order to support learning to prove. In F. Lin, F. Hsieh, G. Hanna, & M. de Villiers (Eds.), *Proceedings of the 19th International Commission on Mathematical Instruction: Proof and Proving in Mathematics Education* (Vol. 2, pp. 130-135). National Taiwan Normal University, Taipei, Taiwan: ICMI Study Series 19, Springer.

Pinto, M., & Tall, D. (1999). Student constructions of formal theory: Giving and extracting meaning. In O. Zaslavsky (Ed.), *Proceedings of the 23rd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 65–73). Haifa, Israel.

Raman, M. (2003). Key ideas: What are they and how can they help us understand how people view proof? *Educational Studies in Mathematics, 52,* 319-325.

Rasmussen, C, & King, K. (2000). Locating starting points in differential equations: A realistic mathematics approach. *International Journal of Mathematical Education in Science and Technology, 31*, 161-172.

Rav, Y. (1999). Why do we prove theorems? *Philosophia Mathematica, 7*, 5-41.

Recio, A. M., & Godino, J. D. (2001). Institutional and personal meanings of mathematical proof. *Educational Studies in Mathematics, 48*, 83-99.

Reid, D. (2002). Conjectures and refutations in grade 5 mathematics. *Journal for Research in Mathematics Education, 33*(1), 5-29.

Preprint

Reid, D. (2005). The meaning of proof in mathematics education. In M. Bosch (Ed.), *Proceedings of the 4th Conference of the European Society for Research in Mathematics Education* (pp. 458-468). Sant Feliu de Guixols, Spain. Retrieved on 11 December, 2001, from http://ermeweb.free.fr/CERME4/CERME4_WG4.pdf

Reid, D. A., & Knipping, C. (2010). *Proof in mathematics education: research, learning, and teaching*. Rotterdam, Boston, Taipei: Sense Publishers.

Reiss, K., Hellmich, F., & Reiss, M. (2002). Reasoning and proof in geometry: Prerequisites of knowledge acquisition in secondary school students. In A. D. Cockburn & E. Nardi (Eds.), *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 113-120). Norwich, England: University of East Anglia.

Remillard, J. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research, 75*(2), 211–246.

Ron, G., & Dreyfus, T. (2004). The use of models in teaching proof by mathematical induction. *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 113-120). Bergen, Norway.

Rowland, T. (2002). Generic proofs in number theory. In S. Campbell & R. Zaskis (Eds.), *Learning and teaching number theory: Research in cognition and instruction* (pp. 157-183). New Jersey: Ablex Publishing Corporation.

Samkoff, A., Lai, Y., & Weber, K. (2012). On the different ways that mathematicians use diagrams in proof construction. *Research in Mathematics Education*, *14*(1), 49-67.

Sandefur, J., Mason, J., Stylianides, G. J., & Watson, A. (2013). Generating and using examples in the proving process. *Educational Studies in Mathematics*, *83*(3), 323-340.

Schoenfeld, A. H. (1980). Teaching problem-solving skills. *American Mathematical Monthly*, 87, 794-805.

Schoenfeld, A. H. (1985). *Mathematical problem solving*. Orlando, FL: Academic Press.

Schoenfeld, A. H. (1991). On mathematics as sense-making: An informal attack on the unfortunate divorce of formal and informal mathematics. In J. F. Voss, D. N. Perkins, & J. W. Segal (Eds.), *Informal reasoning and education* (pp. 311-343). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sears, R. (2012). *An examination of how teachers use curriculum materials for the teaching of proof in high school geometry.* Unpublished doctoral dissertation, University of Missouri, Columbia.

Sears, R., & Chávez, O. (2014). Opportunities to engage with proof: the nature of proof tasks in two geometry textbooks and its influence on enacted lessons. *ZDM – The International Journal on Mathematics Education, 46*, 767-780.

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

Selden, A., & Selden, J. (2013). Proof and problem solving at university level. *Montana Mathematics Enthusiast*, *10*, 303-334.

Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics*, *29*(2), 123-151.

Senk, S. L. (1989). van Hiele levels and achievement in writing geometry proofs. *Journal for Research in Mathematics Education*, *20*, 309-321.

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher*, *27*(2), 4-13.

Preprint

Silver, E. A., Alacaci, C., & Stylianou, D. (2000). Students' performance on extended constructed-response tasks. In P. A. Kenney & E. A. Silver (Eds.), *Results from the seventh mathematics assessment of the National Assessment of Educational Progress* (pp. 301-341). Reston, VA: National Council of Teachers of Mathematics.

Simon, M. A., & Blume, G. W. (1996). Justification in the mathematics classroom: A study of prospective elementary teachers. *Journal of Mathematical Behavior, 15*, 3–31.

Smith, J. C. (2006). A sense-making approach to proof: Strategies of students in traditional and problem-based number theory courses. *Journal of Mathematical Behavior, 25*, 73-90.

Smith, J. P., diSessa A. A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences, 3*(2), 115-163.

Sowder, L., & Harel, G. (2003). Case studies of mathematics majors' proof understanding, construction, and appreciation. *Canadian Journal of Science, Mathematics and Technology Education, 3,* 251-267.

Staples, M., Bartlo, J., & Thanheiser, E. (2012). Justification as a teaching and learning practice: Its multifacted (potential) role in middle school classrooms. *Journal of Mathematical Behavior*, *31*, 447-462.

Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning, 10*(4), 313–340.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom.* New York, NY: The Free Press.

Stylianides, A. J. (2007a). Introducing young children to the role of assumptions in proving. *Mathematical Thinking and Learning, 9*, 361-385.

Stylianides, A. J. (2007b). Proof and proving in school mathematics. *Journal for Research in Mathematics Education, 38,* 289–321.

Stylianides, A. J. (2007c). The notion of proof in the context of elementary school mathematics. *Educational Studies in Mathematics, 65,* 1-20.

Stylianides, A. J. (2009a). Breaking the equation ''empirical argument = proof''. *Mathematics Teaching, 213,* 9–14.

Stylianides, A. J., & Al-Murani, T. (2010). Can a proof and a counterexample coexist? Students' conceptions about the relationship between proof and refutation. *Research in Mathematics Education, 12*(1), 21-36.

Stylianides, A. J., & Ball, D. L. (2008). Understanding and describing mathematical knowledge for teaching: Knowledge about proof for engaging students in the activity of proving. *Journal of Mathematics Teacher Education, 11*, 307-332.

Stylianides, A. J., Bieda, K., & Morselli, F. (forthcoming). Proof and argumentation in mathematics education research. In G. Leder, P. Boero, & A. Gutiérrez (Eds.), *Second handbook of research on the Psychology of Mathematics Education.* Sense Publishers.

Stylianides, A. J., & Stylianides, G. J. (2009a). Proof constructions and evaluations. *Educational Studies in Mathematics*, *72*(2), 237-253.

Stylianides, A. J., & Stylianides, G. J. (2013). Seeking research-grounded solutions to problems of practice: Classroom-based interventions in mathematics education. *ZDM – The International Journal on Mathematics Education, 45*(3), 333-341.

Preprint

Stylianides, A. J., & Stylianides, G. J. (2014a). Impacting positively on students' mathematical problem solving beliefs: An instructional intervention of short duration. *Journal of Mathematical Behavior, 33*, 8-29.

Stylianides, A. J., Stylianides, G. J., & Philippou, G. N. (2002). University students' conceptions of empirical proof and proof by counterexample. In M. Tzekaki (Ed.), *Proceedings of the 5th Panellenian Conference on Didactics of Mathematics and Computers in Education* (pp. 150-158). Volos, Greece: University of Thessaly. (in Greek)

Stylianides, A. J., Stylianides, G. J., & Philippou, G. N. (2004). Undergraduate students' understanding of the contraposition equivalence rule in symbolic and verbal contexts. *Educational Studies in Mathematics, 55*, 133-162.

Stylianides, G. J. (2008a). An analytic framework of reasoning-and-proving. *For the Learning of Mathematics, 28*(1), 9–16.

Stylianides, G. J. (2008b). Proof in school mathematics curriculum: A historical perspective. *Mediterranean Journal for Research in Mathematics Education, 7*(1), 23-50.

Stylianides, G. J. (2008c). Investigating the guidance offered to teachers in curriculum materials: The case of proof in mathematics. *International Journal of Science and Mathematics Education, 6*, 191-215.

Stylianides, G. J. (2009b). Reasoning-and-proving in school mathematics textbooks. *Mathematical Thinking and Learning, 11*, 258-288.

Stylianides, G. J. (2014). Textbook analyses on reasoning-and-proving: Significance and methodological challenges. *International Journal of Educational Research, 64,* 63-70.

Stylianides, G. J., & Stylianides, A. J. (2008). Proof in school mathematics: Insights from psychological research into students' ability for deductive reasoning. *Mathematical Thinking and Learning, 10*(2), 103-133.

Stylianides, G. J., & Stylianides, A. J. (2009b). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education, 40*, 314-352.

Stylianides, G. J., & Stylianides, A. J. (2014b). The role of instructional engineering in reducing the uncertainties of ambitious teaching. *Cognition and Instruction, 32*(4), 1-42.

Stylianides, G. J., Stylianides, A. J., & Philippou, G. N. (2007). Preservice teachers' knowledge of proof by mathematical induction. *Journal of Mathematics Teacher Education, 10,* 145-166.

Stylianides, G. J., Stylianides, A. J., & Shilling-Traina, L. N. (2013). Prospective teachers' challenges in teaching reasoning-and-proving. *International Journal of Science and Mathematics Education 11,* 1463-1490.

Stylianou, D. A., Blanton, M. L., Knuth, E. J. (Eds.). (2010). *Teaching and learning proof across the grades: A K-16 perspective*. Routledge & National Council of Teachers of Mathematics.

Tapan, M. S., & Arslan, C. (2009). Preservice teachers' use of spatio-visual elements and their level of justification dealing with a geometrical construction problem. *US-China Education Review, 6*(3), 54-60.

Tarr, J. E., Chávez, Ó., Reys, R. E., & Reys, B. J. (2006). From the written to the enacted curricula: The intermediary role of middle school mathematics teachers in shaping students' opportunity to learn. *School Science and Mathematics*, *106*(4), 191-201.

Thompson, D. R., Senk, S. L., & Johnson, G. J. (2012). Opportunities to learn reasoning and proof in high school mathematics textbooks. *Journal for Research in Mathematics Education, 43*, 253-295.

Preprint

Tsamir, P., Tirosh, D., Dreyfus, T., Barkai, R., & Tabach, M. (2008). Inservice teachers' judgment of proof in ENT. In O. Figueras, J. L. Cortina, S. Alatorre, T. Rojano, & A. Sepúlveda (Eds.), *Proceedings of the 32nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 345-352). Morelia, Mexico.

Volmink, J. D. (1990). The nature and role of proof in mathematics education. *Pythagoras*, *23*, 7-10.

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics*, *48*(1), 101-119.

Weber, K. (2002). Beyond proving and explaining: Proofs that justify the use of definitions and axiomatic structures and proofs that illustrate technique. *For the Learning of Mathematics*, *22*(3), 14-17.

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. *The Journal of Mathematical Behavior*, *23*(2), 115-133.

Weber, K. (2005). Problem-solving, proving, and learning: The relationship between problem-solving processes and learning opportunities in the activity of proof construction. *The Journal of Mathematical Behavior*, *24*(3), 351-360.

Weber, K. (2006). Investigating and teaching the processes used to construct proofs. *Research in Collegiate Mathematics Education*, *6*, 197-232.

Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education*, *39*, 431-459.

Weber, K. (2010a). Mathematics majors' perceptions of conviction, validity, and proof. *Mathematical Thinking and Learning*, *12*(4), 306-336.

Weber, K. (2010b). Proofs that develop insight. *For the Learning of Mathematics*, *30*(1), 32-36.

Weber, K. (2012). Mathematicians' perspectives on their pedagogical practice with respect to proof. *International Journal of Mathematics Education in Science and Technology*, *43*(4), 463-475.

Weber, K., & Alcock, L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics*, *56*(2-3), 209-234.

Weber, K. & Alcock, L. (2005). Using warranted implications to understand and validate proofs. *For the Learning of Mathematics*, *25*(1), 34-38.

Weber, K. & Alcock, L. (2009). Proof in advanced mathematics classes: Semantic and syntactic reasoning in the representation system of proof. In D. A. Stylianou, M. L. Blanton, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. 323-338). New York: Routledge.

Weber, K., Inglis, M., & Mejía-Ramos, J. P. (2014). How mathematicians obtain conviction: Implications for mathematics instruction and research on epistemic cognition. *Educational Psychologist*, *49*, 36-58.

Weber, K., Maher, C., Powell, A., & Lee, H. S. (2008). Learning opportunities from group discussions: Warrants become the objects of debate. *Educational Studies in Mathematics*, *68*(3), 247-261.

Weber, K., & Mejia-Ramos, J. P. (2009). An alternative framework to evaluate proof productions: A reply to Alcock and Inglis. *The Journal of Mathematical Behavior*, *28*(4), 212-216.

Weber, K. & Mejía-Ramos, J.P. (2011). Why and how mathematicians read proofs: An exploratory study. *Educational Studies in Mathematics, 76*(3), 329-344.

Weber, K., & Mejia-Ramos, J. P. (2013). On the influence of sources in the reading of mathematical text. *Journal of Literacy Research*, *45*, 87-96

Weber, K., & Mejia-Ramos, J.P. (2014). Mathematics majors' beliefs about proof reading. *International Journal of Mathematics Education in Science and Technology*, *45,* 89-103.

Weber, K., & Mejia-Ramos, J. P. (in press). The contextual nature of conviction in mathematics. To appear in *For the Learning of Mathematics.*

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education*, *27,* 458-477.

Yackel, E., & Hanna, G. (2003). Reasoning and proof. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to Principles and Standards for School Mathematics* (pp. 22-44). Reston, VA: National Council of Teachers of Mathematics.

Yang, K. L., & Lin, F. L. (2008). A model of reading comprehension of geometry proof. *Educational Studies in Mathematics*, *67*(1), 59-76.

Yeager, D. S., & Walton, G. M. (2011). Social–psychological interventions in education: They're not magic. *Review of Educational Research, 81*, 267-301.

Yu, J. Y. W., Chin, E. T., & Lin, C. J. (2004). Taiwanese junior high school students' understanding about the validity of conditional statements. *International Journal of Science and Mathematics Education, 2,* 257-285.

Zack, V. (1997). ''You have to prove us wrong'': Proof at the elementary school level. In E. Pehkonen (Ed.), *Proceedings of the 21st Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 291–298). Lahti, Finland: University of Helsinki.

Zandieh, M., Larsen, S., & Nunley, D. (2008). Proving starting from informal notions of symmetry and transformations. In M. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and teaching in undergraduate mathematics education* (pp. 125-138). Washington, DC: Mathematical Association of America.

Zandieh, M., Roh, K. H., & Knapp, J. (2014). Conceptual blending: Student reasoning when proving "conditional implies conditional" statements. *The Journal of Mathematical Behavior*, *33*, 209-229.

Zaslavsky, O. (2005). Seizing the opportunity to create uncertainty in learning mathematics. *Educational Studies in Mathematics, 60*, 297–321.

Zaslavsky, O., Nickerson, S. D., Stylianides, A. J., Kidron, I., & Winicki, G. (2012). The need for proof and proving: Mathematical and pedagogical perspectives. In G. Hanna & M. de Villiers (Eds.), *Proof and proving in mathematics education: The 19th ICMI Study* (pp. 215-229) (New ICMI Study Series, Vol. 15). New York, NY: Springer.

Zazkis, D., Weber, K., & Mejia-Ramos, J. P. (in press). Bridging the gap between graphical argument and acceptable proofs in a university calculus context. To appear in *Educational Studies in Mathematics*.

Zazkis, R., Liljedahl, P., & Chernoff, E. J. (2008). The role of examples in forming and refuting generalizations. *ZDM – The International Journal on Mathematics Education, 40,* 131-141.

Preprint