Reference:

Stylianides, A. J., & Stylianides, G. J. (accepted). Introducing students and prospective teachers to the notion of proof in mathematics. *The Journal of Mathematical Behavior.*

Introducing students and prospective teachers to the notion of proof in mathematics

Andreas J. Stylianides[1] & Gabriel J. Stylianides[2]

Author note

The two authors contributed equally to the preparation of this paper. The paper is based on research supported by funds from the Spencer Foundation to both authors (grant numbers: 200700100, 200800104) and the Economic and Social Research Council (ESRC) in England to the first author (grant reference: RES-000-22-2536). The opinions expressed in the paper are those of the authors and do not necessarily reflect the position, policies, or endorsement of either funding agency. The authors declare no conflicts of interest.

[1] University of Cambridge
Faculty of Education
184 Hills Road
Cambridge CB2 8PQ, UK
E-mail: as899@cam.ac.uk
Tel.: +44 (0) 1223 767550

[2] University of Oxford
Department of Education
15 Norham Gardens
Oxford, OX2 6PY, UK
E-mail: gabriel.stylianides@education.ox.ac.uk

Introducing students and prospective teachers to the notion of proof in mathematics

Abstract: Although the notion of *proof* is important for all learners' mathematical experiences, there has been limited attention to what it might involve and look like to introduce students and prospective teachers to proof. In this paper we argue for the importance of having a coherent approach to introducing students and prospective teachers to proof, and we discuss the theoretical basis of a learning trajectory relevant to both groups. We also discuss an instructional sequence that aimed to promote the learning trajectory among English secondary students and U.S. prospective elementary teachers, drawing on data from two multi-year design experiments. The learning trajectory comprises two milestones: (1) seeing a need to learn about proof and (2) developing an operationally functional conceptualization of proof. The "need" in milestone 1 entails an aspect of epistemological justification applicable to both students and prospective teachers, and a further aspect of pedagogical justification applicable to prospective teachers.

Keywords: Intellectual need; Intervention; Proof; School mathematics; Task design; Teacher education

1. Introduction

In recent decades, many researchers and curriculum frameworks internationally have recommended that *proof* and related notions, such as *argumentation* and *proving*, be part of students' mathematical experiences throughout their schooling and as early as the elementary years (e.g., Department for Education, 2013; NGA & CCSSO, 2010; Norwegian Directorate for Education and Training, 2020; Stylianides, Stylianides, & Weber, 2017). An important reason for these recommendations is the central role that proof can play in students' engagement with mathematics as a sense-making activity, whereby assertions are accepted based on reason and argument rather than by appeal to authority. The wide recognition of the importance of proof for students' mathematical experiences in school signifies that proof should also be part of prospective teachers' mathematical experiences in teacher education, because teachers' knowledge and beliefs about proof shape their readiness, willingness, and capacity to support students' engagement with proof (Bieda, 2010; Buchbinder & McCrone, 2020; Knuth, 2002; Stylianides & Ball, 2008).

Despite wide recognition of the importance of proof for learners' mathematical experiences in both school and mathematics teacher education, the field has paid limited attention thus far to what students and prospective teachers' introduction to proof might involve and look like. Research on this issue is important for at least three interrelated reasons.

First, this research can contribute to elevating the place of proof in school mathematics classrooms. Currently, proof has a marginal place in ordinary mathematics classrooms internationally (e.g., Bieda, 2010; Hiebert et al., 2003; Sears & Chávez, 2014). This marginalization of proof is the result of a synergy of factors related, for example, to teachers' knowledge and beliefs, curricular resources, and testing regimes (Stylianides et al., 2017). Another fundamental obstacle to elevating the place of proof in school mathematics is, we argue, the fact that the field currently lacks ways to introduce students and prospective teachers to the notion of proof that would help them see proof as relevant to and important for their mathematical work as well as serve as a solid foundation for their subsequent engagement with proof.

Second, this research can help begin to redress an imbalance in the current body of proof-related research that has paid much more attention to documenting problems of classroom practice as compared to seeking ways to begin to instructionally address some of these important problems (Stylianides & Stylianides, 2017; Stylianides et al., 2017). Indeed, whereas many studies highlighted weaknesses in students and prospective teachers' knowledge about proof, there is a scarcity of studies or debate on how to introduce students and prospective teachers to the notion of proof, which we view as a necessary (albeit insufficient) step towards their productive engagement with proof. Also, if students and prospective teachers' introduction to proof is consistent with one another, as we aspire it to be, this can facilitate a coherent approach to proof instruction in school and teacher education settings. This coherence is particularly important for prospective teachers as it can support a continuity between their epistemological stance and instructional orientation (Schoenfeld, 1994a) in the area of proof.

Third, this research can complement the literature on researchers' debates about or reports of their own conceptualizations of proof (e.g., Balacheff, 2002; Mariotti, Durand-Guerrier, & Stylianides, 2018; Reid, 2005; Stylianides, 2007a, 2007b; Weber, 2014). When reflecting on the meaning of proof in mathematics education research over 15 years ago, Reid (2005) noted that it was perhaps "a sign of the maturity of research into the teaching and learning of proof and proving that we [were] beginning to reflect on what it is we [were] researching, and whether, as a community, we [were] successful in communicating our work to each other" (p. 1).

We believe that it is time for the field to debate also, and become clearer about, the conceptualization(s) of proof that students and prospective teachers should be introduced to so as to help guide their engagement with proof.[1]

In this paper we take a step towards addressing this need for research by focusing on the following question:

> What might be a learning trajectory for students and prospective teachers' introduction to proof in mathematics, and what might it look like when an instructional sequence aims to promote it?

This question reflects our aim to initiate a discussion in the field about students and prospective teachers' introduction to proof that would be based not only on a theoretical proposition of a learning trajectory but also on a proposition that we would illustrate could be put into practice, in both school and teacher education settings.

While demonstrating the success of the instructional sequence to promote the learning trajectory is beyond the scope of this paper (given space constraints), we deemed illustrating its promise to be an ethical imperative: we did not want to put on others the onus of pursuing the practical implications of our theoretical proposition without us offering some concrete indication that this is indeed a viable endeavor. Also, as it is typical of design-based research (e.g., Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003), the theoretical proposition co-emerged with the instructional sequence that we developed to promote the learning trajectory in two design experiments. Accordingly, reporting only the theoretical proposition without also illustrating its dialectic relationship with the design and implementation of the

[1] The two areas of debate – one concerning researchers' own perspectives on proof and another concerning the perspective(s) on proof to which students and prospective teachers should be introduced – are obviously related and overlapping. Indeed, researchers like us who aspire to develop a conceptualization of proof that students and prospective teachers should be introduced to are inevitably influenced by their own epistemology of proof (Balacheff, 2002).

respective instructional sequence to promote the learning trajectory would fail to justify the value of the theory (Barab & Squire, 2004).

Next we first present our proposed learning trajectory and its theoretical basis. Then we illustrate what it looked like when an instructional sequence that we developed in two design experiments to promote the learning trajectory was implemented with secondary students in England and prospective elementary teachers in the United States. These two groups offer a good variation to illustrate the learning trajectory and the feasibility of promoting it in practice. Towards the end of the paper we discuss possible adaptations to the learning trajectory and respective instructional sequence for use also with elementary students and prospective secondary mathematics teachers.

2. The learning trajectory and its theoretical basis

The learning trajectory is outlined in Fig. 1. Its overarching goal was to introduce students and prospective teachers to the notion of proof in mathematics and it comprised two milestones: (1) seeing a need to learn about proof and (2) developing an operationally functional conceptualization of proof. The "need" in milestone 1 entails an aspect that is common for students and prospective teachers that relates to them seeing an *intellectual need* to learn about proof (milestone 1a), and a further aspect for prospective teachers only that relates to them seeing also a *pedagogical need* to learn about proof (milestone 1b). Before we discuss separately each milestone, we make six general comments about the learning trajectory.

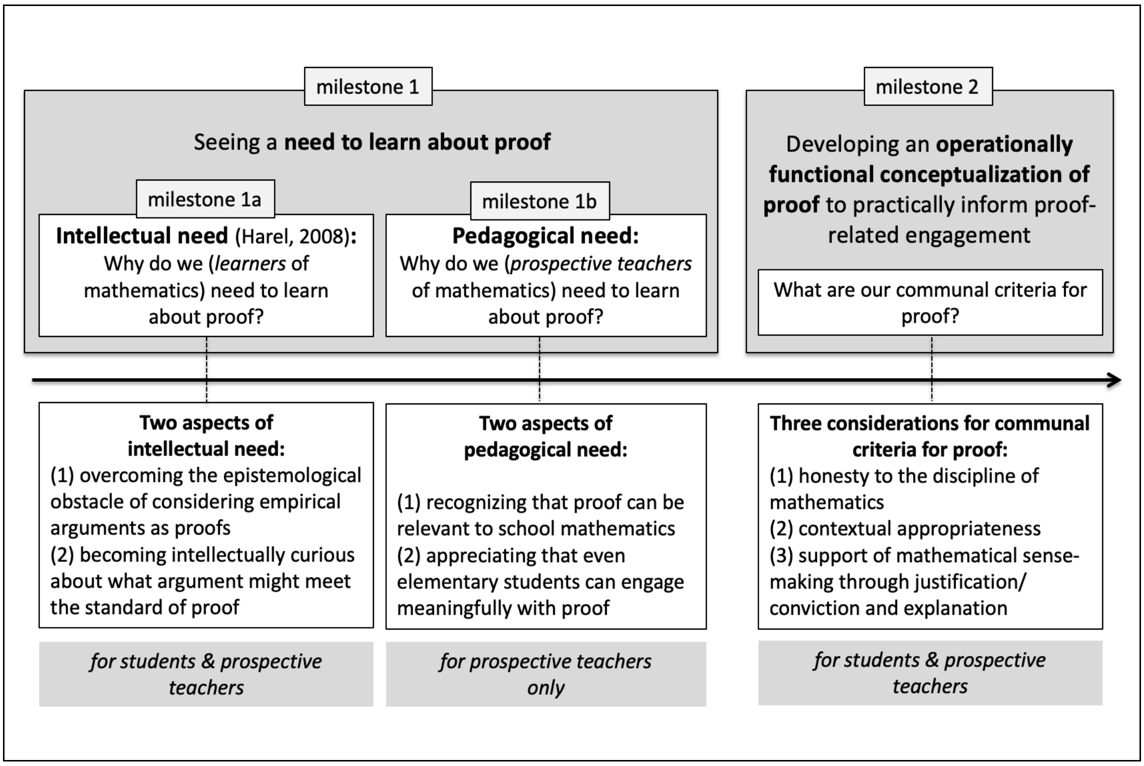

Fig 1 An outline learning trajectory for introducing students and prospective teachers to the notion of proof in mathematics

First, learners' introduction to proof can be a long process comprising many milestones. The scope of our proposed learning trajectory is modest: It comprises a small number of milestones that we view as important when considering students and prospective teachers' *introduction* to the notion of proof and that can be promoted using an instructional sequence of a relatively short duration (extending over a couple of lessons). We emphasize "introduction" to indicate that we are not intending this learning trajectory to be the endpoint of students and prospective teachers' engagement with proof. Rather, we view the learning trajectory as marking the beginning of their proof-related work, preparing them for a productive subsequent engagement with proof. Having such a learning trajectory is an acknowledgment both of the fact that an introduction to proof deserves explicit instructional attention and of the reality of content-packed school and teacher education mathematics curricula that would make difficult the incorporation in them of a more extended introduction.

Second, we view the final milestone of learners' introduction to proof in the learning trajectory as being the development of an *operationally functional* conceptualization of proof, that is, a set of criteria that classroom communities of students and prospective teachers, or individual members of these communities, can use to practically inform their engagement with proof. This, however, does not mean that the conceptualization of proof will allow, unambiguously and from the start, the members of a classroom community to distinguish between arguments that meet the standard of proof and others that do not. The expectation is that the meaning of the communal criteria for proof will be clarified over time and negotiated through social interactions amongst the members of a classroom community including the instructor.

Third, an obstacle to the development of communal criteria for proof is that many students and prospective teachers do not see a *need* to learn about proof from an epistemological standpoint (Harel, 2008; Stylianides & Stylianides, 2009b), while many prospective teachers also fail to see a need to learn about proof from a pedagogical standpoint, viewing proof as an "advanced topic" that is appropriate only for a select group of students (Bieda, 2010; Knuth, 2002). Indeed, a distinguishing feature of our learning trajectory compared to other ways of introducing learners to communal criteria for proof (Campbell & King, 2020; Yee, Boyle, Ko, & Bleiler-Baxter, 2018) is its provision for students and prospective teachers to see a need to learn about proof (in the sense outlined in milestone 1) *before* they are introduced to an operationally functional conceptualization of proof (milestone 2).

Fourth, although the learning trajectory as a whole is a unique contribution to the mathematics education literature in the area of proof, prior research has informed its development as we will discuss later. Also, a small number of prior studies addressed in isolation ideas related to milestones 1 or 2. Regarding milestone 2, a few studies examined school and university students' use of communal criteria for proof in their argument constructions or evaluations (Campbell & King, 2020; Stylianides, 2019; Stylianides & Stylianides, 2009a; Yee et al., 2018); yet how to introduce students and prospective teachers, in actual classroom practice, to an operationally functional conceptualization of proof was not an explicit concern of these studies. Regarding milestone 1, although some research considered ways of creating among students and prospective teachers a need to learn about proof from an epistemological standpoint (Brown, 2014; Harel, 2008; Stylianides & Stylianides, 2009b, 2014b), there has been no research on how to help prospective teachers see a need to learn about proof from a pedagogical standpoint. The bottom line is that prior research in areas relevant to the learning trajectory has not been integrated into a single coherent learning trajectory nor has a way of promoting such a learning trajectory been illustrated in its entirety before.

Fifth, we view the construct of a *learning trajectory* as referring to a goal (in this case, to introduce students and prospective teachers to the notion of proof in mathematics) and a learning progression (reflected in the two milestones we described earlier), as summarized in Fig. 1. These features capture two of the three features that Clements, Sarama, Baroody, and Joswick (2020) proposed should characterize a learning trajectory. We agree with Clements et al. that it is vitally important for educationalists to design *instructional sequences* to promote their specified goals and learning progressions, but, contrary to them and consistent with the conceptualization of a learning trajectory that we used in our prior work (e.g., Stylianides & Stylianides, 2009b, 2014b), we do not view the instructional sequences as part of the notion of the learning trajectory. For us, the term "*learning* trajectory" draws attention more to the learner–content side of the "instructional triangle" (Cohen, Raudenbush, & Ball, 2003), and so we consider useful to label separately the instructional sequences that aim, through the teacher, to enable the desirable learner–content relationship.

Sixth, the particular learning trajectory can be "hypothetical" (Simon, 1995), representing the learning path students and prospective teachers can be expected to follow during the implementation of an instructional sequence that aims to promote the learning trajectory, or "actual" (Leikin & Dinur, 2003), representing the learning path students and prospective teachers actually followed during the implementation of such an instructional sequence. Also, the trajectory can be "individual," representing the learning path of individual students or prospective teachers in a school or a teacher education class, or "communal," representing the learning path of a classroom community as a whole, which might differ from the learning paths of individual classroom participants (Stylianides & Stylianides, 2009b, 2014b). Even the existence of "communal criteria for proof" (milestone 2) can be understood or used variably by individual classroom participants. Our presentation of the particular learning trajectory is not contingent upon it being viewed as hypothetical or actual, individual or communal.

2.1. A need to learn about proof (milestone 1)

According to Harel's (2008) necessity principle, "[f]or students to learn about the mathematics we intend to teach them, they must have a need for it" (p. 900). Harel specified this need in terms of "intellectual need," which corresponds to milestone 1a and applies to both students and prospective teachers. Milestone 1b refers to another kind of need, which we call "pedagogical need," relevant only to prospective teachers.

2.1.1. Intellectual need (milestone 1a)

The notion of intellectual need is "inextricably linked to the notion of epistemological justification" (Harel, 2013, p. 120) and "has to do with disciplinary knowledge being born out of people's current knowledge through engagement in problematic situations conceived as such by them" (Harel, 2008, p. 898). According to Harel (2008, 2013), learners are often epistemologically resistant to changing some of their current conceptions, despite the conceptions deviating from conventional mathematical understandings, and this is because instructors fail to present learners with a clear intellectual purpose for what the instructors want to teach.

In the particular area of proof, substantial research evidence shows that learners of all ages, including prospective teachers, tend to hold views about the meaning of proof that deviate from broadly accepted meanings (for reviews see: Harel & Sowder, 2007; Stylianides et al., 2017). Accordingly, learners' perceived proofs frequently fall short of satisfying a mathematical standard of proof and thus are *non-proof arguments* (Stylianides, 2008). *Empirical arguments*, that is, invalid arguments that purport to establish the truth of a mathematical generalization based on the confirming evidence obtained from examining a proper subset of all the possible cases covered by the generalization, are arguably the most notable kind of non-proof argument perceived by learners to be a proof (Education Committee of the European Mathematical Society, 2011; Harel & Sowder, 2007; Stylianides et al., 2017).

According to the Education Committee of the European Mathematical Society (2011), a "solid finding" of mathematics education research is that many learners "rely on validation by means of one or several examples to support general statements, that this phenomenon is *persistent* in the sense that many [learners] continue to do so even after explicit instruction about the nature of mathematical proof, and that the phenomenon is *international*" (pp. 50-51; emphasis added). Indeed, no other phenomenon in the area of proof appears to have attracted so much attention by mathematics education researchers. Also indicative of the important epistemological place of empirical arguments in learners' progression to proof is that all major classifications of arguments that learners consider to be proofs (reviewed in Stylianides et al., 2017) include at least one category of empirical arguments immediately prior to the categories of arguments that meet the standard of proof. In one of the most influential hierarchies of arguments by Balacheff (1988), two categories of empirical arguments ("naïve empiricism" and "crucial experiment") is all there is prior to the proof categories ("generic example" and "thought experiment"). In another influential framework, Harel and Sowder's (2007) proof schemes, the empirical category ("empirical proof schemes") again immediately precedes the proof category ("deductive" or "analytic proof schemes").

Therefore, addressing the epistemological obstacle of considering empirical arguments as proofs is an important prerequisite for instruction in order to create an intellectual need among students or prospective teachers to learn about proof (Brown, 2014; Harel, 2008; Stylianides & Stylianides, 2009b) (aspect 1 of intellectual need, Fig. 1). We emphasize that the focus here is on what counts as a proof in the learners' eyes, for many of whom empirical arguments constitute conclusive acts of validation, and should not be misinterpreted as a lack of appreciation of empirical explorations in proving; indeed, examples have important roles and uses in the proving process, including in helping formulate conjectures and offer insights into the development of proofs (Stylianides, 2008; Zaslavsky, Knuth, & Ellis, 2019). Once learners begin to view proof as a class of non-empirical arguments, they can be expected to become intellectually curious about what kind of a non-empirical argument might meet the standard of proof (aspect 2 of intellectual need). This is not to deny the fact that even mathematicians in some cases see value in numerical validations (Bailey & Borwein, 2005) or gain conviction by deference to authority or other non-deductive means (Weber, Inglis, & Mejía-Ramos, 2014); rather it is to emphasize the importance of learners understanding the distinctions between empirical and non-empirical validations and the prominent role of the latter in mathematical practice in the area of proof (Brown, 2014; Harel & Sowder, 2007; Stylianides, 2007a).

2.1.2. Pedagogical need (milestone 1b)

Although having an intellectual curiosity can mark students' readiness to be introduced to an operationally functional conceptualization of proof, the situation is different for prospective teachers. Prospective teachers need to see further a *pedagogical need* to learn about proof, which we define as the need for them to learn about proof from a teacher's standpoint.

Prospective teachers tend to resist learning mathematics that they perceive have little relevance to teaching (Stylianides & Stylianides, 2014a; Wasserman, Weber, Fukawa-Connelly, & McGuffey, 2019). Regarding proof in particular, prior research showed that some secondary mathematics teachers view proof as an "advanced topic" and engagement with proof as a goal appropriate only for a select group of students (Bieda, 2010; Knuth, 2002). Given these findings, one can reasonably expect that many elementary teachers, who teach younger students and whose experiences with proof were likely limited to their own secondary schooling, would view proof as a notion that is beyond the reach of elementary students. All of these highlight the importance of helping prospective teachers, especially elementary teachers, to recognize that proof can be relevant to school mathematics (aspect 1 of pedagogical need, Fig. 1) and appreciate that even elementary students can engage meaningfully with proof (aspect 2 of pedagogical need).

It is important to note that the aim here is not for prospective teachers to see a pedagogical need to learn about *any* conceptualization of proof but rather a need to learn about a conceptualization of proof that is compatible with the one to which prospective teachers will be introduced in milestone 2. Consider, for example, a group of prospective elementary teachers holding the encompassing conception that "proof is any sort of explanation" or a group of prospective secondary mathematics teachers whose conceptualization of proof is restricted to two-column proofs in geometry (cf. Herbst, 2002). Neither of these conceptualizations present defensible ways of thinking about the notion of proof in school mathematics and beyond (Stylianides, 2007a, 2016), and so both groups of prospective teachers would need to see a pedagogical need to learn about proof in a new way, before they could be ready to be introduced to the operationally functional conceptualization of proof that we discuss next.

2.2. An operationally functional conceptualization of proof (milestone 2)

Learners' introduction to proof would be lacking a necessary foundation without a set of criteria that could serve as points of reference to guide their proof-related work (e.g., "What sort of argument should we aim for when we want to develop a proof?") or as a framework to allow learners to reflect on or evaluate this work (e.g., "Should we count this argument as a proof? If not, how might we modify or improve it?") (e.g., Campbell & King, 2020; Yee et al., 2018). This is not to say, however, that once given or presented with proof criteria, the members of a classroom community will interpret each criterion the same way or as intended by the instructor. Indeed, as we will discuss next, even mathematics education researchers have different perspectives on the meaning of some key terms like "explanation" that could be used potentially in the formulation of proof criteria.[2] No matter how clearly defined the proof criteria are, the meaning of these criteria will still have to be socially negotiated within a classroom community – both at the point of the introduction of the criteria and over time during the application of the criteria in proving tasks – so as to progressively support the emergence of a shared understanding amongst its members (Mariotti, 2006; Stylianides, 2007b; Yackel & Cobb, 1996).

[2] There is no agreement among researchers on a set of proof criteria in the first place.

Thus, our position on learners' development of an operationally functional conceptualization of proof (milestone 2) is that, once learners see a need to learn about proof (milestone 1), they should be introduced to proof criteria that meet certain *considerations*, even though it is to be expected that learners will interpret and use these criteria in different ways, at least at the beginning. These criteria will constitute the basis and starting point of learners' subsequent proof-related work, and their formulation will have to take into account prior research knowledge on the *meaning* of proof and key *functions* (i.e., purposes or goals) proofs can serve in learners' mathematical work. Regarding considerations to guide the selection or formulation of proof criteria, we propose the following (Fig. 1): (1) honesty to the discipline of mathematics, (2) contextual appropriateness, and (3) support of mathematical sense-making through justification/conviction and explanation. Consideration 2 implies that certain criteria cannot apply universally across different settings, while considerations 1 and 3 imply that contextually modified criteria should still abide by some core principles.

2.2.1. Meaning of proof

Discussions or debates about the meaning of proof in mathematics education research tended to revolve around theoretical issues and methodological ramifications of researchers' use of unclear, variant, or inconsistent meanings of proof (e.g., Balacheff, 2002; Mariotti et al., 2018; Reid, 2005). In recent years researchers have started to address also how existing conceptualizations of proof can be used in the classroom at the school, university, or teacher education levels (Campbell & King, 2020; Stylianides, 2019; Stylianides & Stylianides, 2009a; Yee et al., 2018), though the process of introducing learners to proof criteria was not the focus of their publications.

The latter group of researchers used as a theoretical basis for their proof criteria the definition of proof that was proposed by Stylianides (2007a); we do the same in this paper. According to Stylianides (2007a), a *proof* in the context of a classroom community at a given time is an argument for the truth or falsity of a mathematical statement that has the following three 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, 2007a, p. 291; emphases in original)

In his elaboration on the definition, Stylianides (2007a, 2007b, 2016) explained that the definition seeks to achieve a balance between honesty to the discipline of mathematics and contextual appropriateness. Regarding honesty to the discipline of mathematics (consideration 1 under milestone 2, Fig. 1), the definition respects the mathematical integrity of a proof by imposing certain requirements on arguments that meet the standard of proof: these arguments need to use true statements, valid modes of argumentation, and appropriate modes of representation, whereby the terms "true," "valid," and "appropriate" should be understood in relation to what is typically agreed upon nowadays in the field of mathematics, in the context of specific mathematical theories. Regarding contextual appropriateness (consideration 2), the definition supports a rather "elastic" meaning of proof that takes account of the current state of knowledge or conceptual capabilities of a particular classroom community at any level of education. On the basis of consideration 1 empirical arguments are disqualified from the class of proofs due to their use of invalid modes of argumentation. Empirical arguments are also disqualified from the class of proofs on the basis of consideration 2: findings from both psychological and mathematics education research show that non-empirical modes of argumentation are within even young students' conceptual reach (for a review of relevant literature, see: Stylianides & Stylianides, 2008). Thus, learners who have not gone through milestone 1a and view empirical arguments as proofs are unlikely to interpret as intended a proof criterion that (directly or indirectly) specifies proof as a non-empirical argument (supportive evidence for this proposition is found in studies that omitted milestone 1a: Campbell & King, 2020; Yee et al., 2018).

The two considerations find philosophical support in the works of general education scholars (e.g., Bruner, 1960; Schwab, 1978) who argued for subject areas being represented in the educational process in ways that are both honest to their respective disciplines (consideration 1) and honoring of learners' current state of knowledge (consideration 2). Further support for the two considerations can be found in the works of mathematics education researchers in the area of proof. According to Mariotti (2006), "the crucial point that has emerged from different research contributions [in the area of proof] concerns the need for proof to be acceptable from a mathematical point of view [consideration 1] but also to make sense for students [consideration 2]" (p. 198).

2.2.2. Functions of proof

An operationally functional conceptualization of proof cannot be viewed in isolation from the functions that proof can serve in mathematical work. Yet there are many important functions a proof can play in learners or mathematicians' work – including explanation, verification/falsification, communication, generation of new knowledge, and systematization (Bell, 1976; de Villiers, 1990; Hanna, 1990; Stylianides, 2009b) – and so the question arises as to which of these functions to highlight when *introducing* learners to proof. In grappling with this question, we found useful Harel and Sowder's (2007) account of learners' *engagement with* *mathematics as a sense-making activity* (consideration 3 under milestone 2, Fig. 1). According to Harel and Sowder (2007), "[m]athematics as sense-making means that one should not only ascertain oneself that the particular topic/procedure makes sense, but also that one should be able to *convince* others through *explanation* and *justification* of her or his conclusions" (pp. 808-809; emphases added).

Thus, one primary function of proof emerging from Harel and Sowder's notion of mathematical sense-making is *justification*, which can be broadly defined as showing *that* a mathematical statement is true (Bell, 1976). The notion of justification is interlinked with the idea of proof as a means for promoting *conviction*, which can take different forms as in Mason's (1982) tripartite hierarchy "convince yourself, convince a friend, convince a skeptic." Conviction in the field of mathematics is often linked to the feeling of certainty that a proof can yield to mathematicians that a statement "*has to be* true" (Schoenfeld, 1994b, p. 74; italics in original), though there is disagreement as to whether mathematicians prove to gain (absolute) certainty (Weber et al., 2014). An indication of the importance of conviction as a proof function is that *proving as convincing* has been identified as one of three main perspectives from which mathematics education researchers have investigated the activity of proving (Stylianides et al., 2017).[3]

The second primary proof function that emerges from Harel and Sowder's notion of mathematical sense-making is *explanation*, which can be broadly defined as offering insight into *why* a mathematical statement is true (Bell, 1976; de Villiers, 1990; Hanna, 1990). Although there is no agreement among mathematics education researchers as to the characterization of explanatory arguments (Lockwood, Caughman, & Weber, 2020; Stylianides, Sandefur, & Watson, 2016), many

[3] The other two perspectives discussed in Stylianides et al. (2017) were *proving as problem solving* and *proving as a socially embedded activity*.

researchers endorse the idea that explanation is an important function of proof in learners' mathematical work (Hanna, 1990; Stylianides et al., 2016; Weber, 2010) with some of them even describing proofs as a special class of "explanatory arguments" (Knuth, 2002) or as a sub-class of "explanations" (Balacheff, 1988).

While it is important to acknowledge justification/conviction and explanation as important proof functions in their own right, it is perhaps unnecessary for the purposes of learners' introduction to proof to try to make clear distinctions between these functions or between "proofs that explain" and "proofs that only convince" (Hanna, 1990; Lockwood et al., 2020). According to Mariotti (2006), explanation and conviction are interlinked as students engage with and develop ownership of the proving activity in a social setting: "students *explain* their arguments to a peer or to the whole class, including the teacher, also to *convince* themselves [or others] of their truth" (p. 198; italics added). Thus, as far as proof functions are concerned, when formulating proof criteria for learners' mathematical work it might suffice to try to capture Schoenfeld's (1994b) statement that "[w]hen you have a proof of something you know it *has to be* true, and *why*" (p. 74; italics added to "why").

3. Two design experiments

To illustrate what it might look like when an instructional sequence aims to promote the learning trajectory, we draw on data from two *design experiments* (Cobb et al., 2003) where we designed one such instructional sequence and implemented it with prospective elementary teachers in the United States and secondary students in England. These two groups offer a good variation to illustrate the learning trajectory and the feasibility of promoting it in practice in different contexts. As we noted earlier, we aimed for the learning trajectory for introducing students and prospective teachers to be as consistent as possible between the two groups; we aimed for the same level of consistency in designing the instructional sequence so as to facilitate a coherent approach to proof instruction in school and teacher education settings.

Full documentation of the implementation of the instructional sequence in either setting is beyond the scope of this paper. Our aim is to strike a balance between describing the theoretical basis of the instructional sequence, which co-emerged with the design and refinement of the instructional sequence over the cycles of our design-based research, and illustrating how the instructional sequence played out during its implementation. Accordingly, there is no suggestion that the sequence's implementation represented a fully successful way to promote the learning trajectory, only a promising one.

3.1. Research context

In line with the tenets of design experiment methodology, the two studies involved theorization and iterative empirical testing and refinement of instructional sequences that aimed to facilitate learner progression along particular learning trajectories (Stylianides & Stylianides, 2014b). Although each study addressed a wide range of topics in various mathematical domains (algebra, geometry, etc.) as these were relevant to the respective research participants, both studies treated proof as a vehicle to mathematical sense-making across domains. Accordingly, some of the learning trajectories or parts thereof, like milestones 1a and 2 of the focal learning trajectory, were relevant to the participants of both studies. The respective instructional sequence was the first one that we implemented in the participating classes in both studies, and this aligned with the purpose of the learning trajectory to mark the beginning of participants' proof-related work and prepare them for a productive engagement with proof.

The first design experiment was a 4-year study that we conducted in an undergraduate mathematics course for prospective elementary teachers in the United States. In the last research cycle of the study, from which the data for this paper derive, the course included two parallel classes that followed the same curriculum, were taught by the second author, and included a total of 39 students. The course was prerequisite for admission to a masters-level elementary teacher education program, and the students were primarily third year undergraduates majoring in different fields of study. For many students this was their first mathematics course since secondary school, and they tended to have weak mathematical backgrounds. Although it was beyond the scope of the course to teach prospective teachers about mathematics pedagogy (this was done in a follow up course), the course did pay attention to the application or relevance of the targeted learning goals to the work of mathematics teaching (Ball, Thames, & Phelps, 2008; Stylianides & Stylianides, 2014a). This implied that a learning trajectory in the course for introducing prospective teachers to the notion of proof in mathematics would also need to consider the creation of a pedagogical need for prospective teachers to learn about proof (milestone 1b).

The second design experiment was conducted by the first author in two secondary mathematics classes in a state school in England. The classes were taught by their regular teachers over a 2-year period, when the students were between 14 and 16 years old. All 61 students from the two highest attaining classes in their year group (out of a total of seven classes) participated in the study. The focus on high-attaining students was guided by the participating teachers' wish to work with this particular group of students in the project and was deemed reasonable given the findings of a previous large-scale longitudinal study in England (Küchemann & Hoyles, 2001-03) that raised concerns about English high-attaining students' knowledge about proof and their readiness for advanced mathematical courses. The study involved the design, implementation, and analysis of six instructional sequences that related to proving and were embedded in a range of topics according to the national mathematics curriculum; each instructional sequence lasted between one and five 45-minute lesson periods. At the beginning of the study, the researcher took main responsibility for planning the instructional sequences, though all plans were subject to discussion with and modification by the teachers prior to their implementation. Over time, the teachers took more responsibility for planning the sequences. The planning for the focal instructional sequence was done by the researcher and was an adapted version of the intervention we had designed and iteratively refined previously in the university study.

The data sources most relevant to this paper are: transcripts produced based on video and audio records of the implementation of the focal instructional sequence with prospective teachers and secondary students; field notes taken by a research assistant; and copies of participants' classwork, including participants' written responses to specific prompts (what we call "conceptual awareness pillars", see below for elaboration) at strategically selected points during the intervention that gave us "snapshots" of participants' individual state of understanding with respect to the learning milestones. In our discussion of milestone 1b, we also draw on participants' responses to three multiple-choice questions in a pre- and post-course survey that we administered in the university study to explore prospective teachers' perceptions about various aspects of mathematical practices and mathematics teaching; we adapted several of the questions in the survey from Schoenfeld (1989), and these three were the only questions relevant to this paper.

3.2. General features of the instructional approach in the two studies

The instructional sequences in the two studies shared some common design features but also had some unique features that were tailored to their specific aims and respective learning trajectories. For this reason, we discuss our instructional approach to promoting each milestone of the focal learning trajectory separately in section 4. In this section we briefly present four general features of our instructional approach in the two studies that applied also to the focal instructional sequence.

First, we aimed for the instructional sequences to have a rather narrow and well-defined scope, which in turn could allow the instructional sequences to have also a relatively short duration and standalone nature (Stylianides & Stylianides, 2017). The modest scope of the focal learning trajectory, discussed earlier, aligns with this instructional design feature.

Second, each instructional sequence comprised instructional materials – notably tasks and prompts – and a description of instructor actions in implementing those materials in the classroom. The tasks comprised different kinds of mathematics tasks, such as tasks with emerging true and failing patterns (as in the part of the instructional sequence addressing milestone 1a), and mathematics tasks contextualized in a pedagogical context (as in the part addressing milestone 1b). The prompts took primarily the form of "conceptual awareness pillars" that we discuss in the third feature below. The tasks and prompts were carefully designed and sequenced based on our hypothesized relationships between the instructional materials and participants' learning progressions, facilitated by the instructor's actions that we discuss in the fourth feature below. Through "empirical tinkering" (Morris & Hiebert, 2011) from one research cycle to the next, and based on our evolving theoretical understanding of the relationships at play between instructional design and learning, we made progressive improvements to the instructional sequences to better support the respective learning trajectories.

Third, a key role in many instructional sequences was played by a special type of prompts that we call *conceptual awareness pillars* or simply *pillars* (Stylianides & Stylianides, 2009b). This is a notion that emerged from our design-based research, and we use it to describe instructional activities that aim to direct students or prospective teachers' attention to their conceptions (understandings, beliefs, pedagogical dispositions, etc.) about a particular mathematical topic or issue. As we explain later, we used pillars in two ways in the focal instructional sequence: (1) to trigger and support the resolution of cognitive conflicts among learners in helping them overcome the epistemological obstacle of considering empirical arguments as proofs (milestone 1a); and (2) to direct prospective teachers' attention to the role and meaning of proof in school mathematics in helping them see a pedagogical need to learn about proof (milestone 1b). The extent to which the pillars fulfilled their intended purposes was examined over the cycles of our design-based research, which supported incremental improvements to the phrasing and sequencing of the pillars.

Fourth, the instructor played a critical role in the implementation of the instructional sequences and was viewed as the representative of the mathematical community in the classroom (Mariotti, 2006; Stylianides & Stylianides, 2009b; Yackel & Cobb, 1996) and, in the university study, of the teacher education community too. The instructor supported the classroom participants to develop knowledge that was consistent with conventional understandings through social interactions around the planned activities (Yackel & Cobb, 1996). The outcomes of these interactions became part of the evolving "taken-as-shared" knowledge of the class (Simon & Blume, 1996; Yackel & Cobb, 1996) through "situations for institutionalization" (Brousseau, 1981). The instructor orchestrated situations for institutionalization and other whole class discussions by implementing the following *four key practices* (for elaboration, see Stylianides & Stylianides, 2014b): (1) soliciting multiple contributions so that different voices could be heard and various ideas could be considered and discussed; (2) asking participants to explain their thinking for their contributions; (3) inviting participants to listen and respond to others' contributions; and (4) revoicing or highlighting selected contributions in order, for example, to help direct the classroom community's attention to ideas relating to specific milestones in the respective learning trajectory. A well-designed set of tasks and pillars, complemented by the instructor's use of practice 4, supported the emergence of patterns in learners' contributions during whole class discussions that increasingly (over the cycles of our design-based research) clustered around the intended learning milestones.

4. The instructional sequence

4.1. Promoting milestone 1a: intellectual need

As we explained earlier, creating an intellectual need for proof involves removal of learners' epistemological obstacle of considering empirical arguments as proofs. This includes empirical arguments of all kinds including Balacheff's (1988) "naïve empiricism" and "crucial experiment." In *naïve empiricism* the examined cases in the generalization under consideration are selected for no particular reason or on the basis of practical convenience, while in *crucial experiment* the examined cases are selected according to a rationale such as a strategy for discovering possible counterexamples.

According to Harel (2008, 2013), the removal of an epistemological obstacle is best experienced through a problematic situation that helps an individual or a community become aware of the limitations of their current knowledge (thus experiencing a state of disequilibrium) and prompts them to develop new knowledge to resolve the problematic situation (thus reaching a new state of equilibrium). To remove learners' epistemological obstacle of considering empirical arguments as proofs, we used *cognitive conflict* (Piaget, 1985; see also Zaslavsky, 2005) as a mechanism to support a series of disequilibrium-equilibrium phases from a naïve empirical conception of proof to a crucial experiment conception of proof and, subsequently, to a non-empirical conception of proof. To evoke and facilitate the resolution of the emerging cognitive conflicts and respective stepwise progressions in learners' conceptions about proof, we used a sequence of three tasks and four conceptual awareness pillars (CAPs) as outlined in Fig. 2. The sequence was almost identical in the two studies with only a slight modification in Task 1, as we explain below. We only give details of Task 1 in our description of how we promoted milestone 1a due to space constraints and because Task 1 was used again for the promotion of milestone 1b in the intervention. Detailed report of each implementation goes beyond the scope of this paper and can be found in Stylianides (2009a) and Stylianides and Stylianides (2009b) for the school and university studies, respectively.

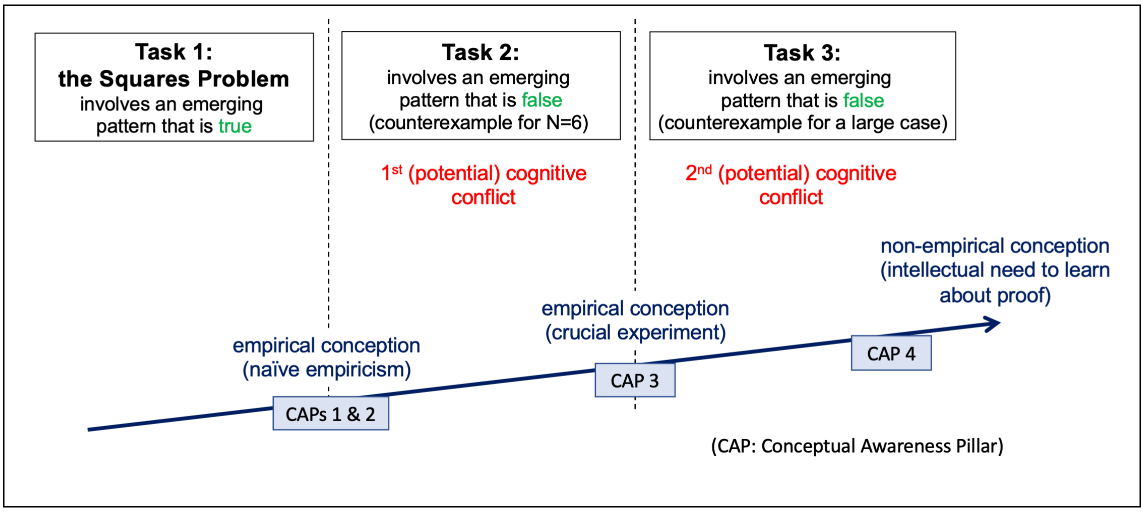

Fig 2 Outline of the part of the instructional sequence used in both studies for milestone 1a

Task 1 was the Squares Problem that we adapted from Zack (1997) and used in slightly different but mathematically similar forms in the two studies (Fig. 3).[4] This involved an emerging pattern that was true and lent itself to empirical validation. As expected from prior research and our experience in previous research cycles, the participants in both studies identified the emerging pattern in their work on the first two questions of the problem for $N = 4$ and 5, and, on the basis of this confirming evidence, they accepted the pattern as true and applied it for N=60 to respond to the third question thus demonstrating a naïve empirical conception.

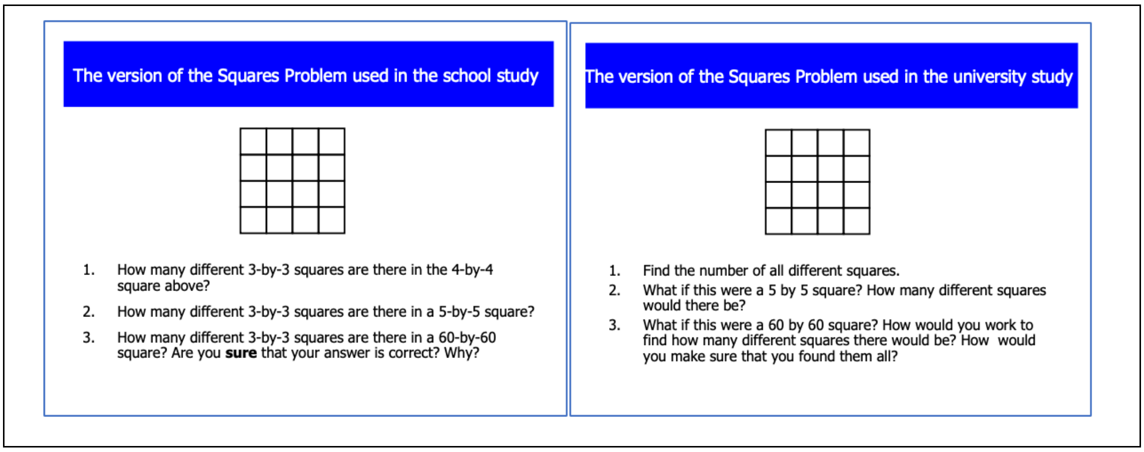

Fig 3 The versions of the Squares Problem (adapted from Zack, 1997) used in the two studies

To help participants become more aware of their validation method, we used the first conceptual awareness pillar (CAP 1). This was a prompt asking participants to write individually whether they were sure that their answer for N=60 was correct, and why. Dan's[5] response from the school study illustrates the general trend in participants' thinking within the realm of naïve empiricism: "I am sure that the answer [for N=60] is correct because it has been proved for a number of smaller grids [for $N = 4$ and 5]."

[4] We used a slightly simpler version of Task 1 with the school students so as to address the collaborating teachers' request for the sequence to have a shorter duration.
[5] We use pseudonyms for all research participants we refer to in the paper.

The other two tasks in the sequence had embedded in them a "plausible pattern" (Stylianides, 2008) that failed for N=6 in Task 2 and for a large case (in the order of septillions) in Task 3. Our choice and sequencing of these tasks and respective counterexamples were deliberate, theoretically informed by the notions of "example spaces" (Watson & Mason, 2005) and "pivotal counterexamples" (Zazkis & Chernoff, 2008), so as to trigger two cognitive conflicts for learners and challenge sequentially their naïve empirical and crucial experiment conceptions. The CAPs before each potential cognitive conflict directed learners' attention to their current conceptions thus helping them become more aware of these conceptions. The CAPs after each contradictory situation (which had the form of an unexpected counterexample to an empirically validated and apparently true pattern) directed learners' attention to the implications of the contradictory situation for their previously expressed conceptions thus facilitating a process of reflection, experience of a cognitive conflict, and subsequent modification of these conceptions.

Dan's written response to CAP 3, which followed Task 2, illustrates participants' transition from naïve empiricism towards a more strategic selection of cases when validating mathematical generalizations, as in crucial experiment. Specifically, Dan wrote that the experience of working on Task 2 taught him he could not always trust a formula that worked for the first few cases (cf. naïve empiricism); rather, he would have to try "spread cases" to test patterns (cf. crucial experiment). By the end of Task 3, participants' faith in empirical arguments of any kind, including crucial experiment, was shaken, with several participants expressing the view in the context of CAP 4 that one had to check all cases in a generalization so as to avoid missing a "hidden" counterexample when validating the generalization. The students recognized, however, that checking all cases in a generalization is not always possible. When the secondary teacher asked, "When do you trust a pattern then?"

Adam said: “When you cannot find one [a counterexample], until you are dead!” Adam’s and other students’ disappointment in the “absence” (from their point of view) of a feasible and secure method to validate mathematical generalizations marked the emergence of an intellectual need for the class to be introduced to proof as such a method.

The classroom discussions and outcomes of implementing the instructional sequence in the school and university classes to promote milestone 1a were markedly similar to each other, as we reported in Stylianides and Stylianides (2014b). Further supportive evidence for the promise of our instructional sequence to promote milestone 1a has been provided by subsequent publications of other researchers. For example, Brown (2014) used similar tasks with undergraduate students and reported that, once the students saw an intellectual need for a non-empirical proof conception, the students did not revert to previous empirical validation practices. Also, Gal (2019) identified students’ lack of cognitive preparedness as a reason for teachers’ failure to evoke a cognitive conflict among students, which reinforces our staged approach to supporting changes to learners’ proof conceptions and our use of CAPs prior to each potential cognitive conflict.

4.2. Promoting milestone 1b: pedagogical need

As indicated by prospective teachers’ responses to the three relevant questions of a survey that we administered at the beginning of cycle 5 of the university study (Table 1), many prospective teachers began the course with a restricted view of proof as an argument presented in two columns in the context only of geometry (items 1 and 2). This was unsurprising because many students and adults in the United States would have first and only encountered proof in its “two-column format” in high school geometry courses (Herbst, 2002); this format emphasized getting proof in the “proper” form rather than mathematical sense making (Schoenfeld, 1989). These experiences created an obstacle to our prospective teachers seeing the relevance of proof to school mathematics as an argument not restricted to specific representational forms or mathematical domains (cf. aspect 1 of pedagogical need in Fig. 1). Relatedly, at the beginning of the course many prospective teachers considered proof to be an advanced topic beyond the scope of the elementary school (item 3). This view created a further obstacle to them appreciating that elementary students can engage meaningfully with proof (cf. aspect 2 of pedagogical need). As indicated by the results of the comparison between prospective teachers’ pre- and post-course responses to the three survey questions (Table 1), the prospective teachers, according to their self-reports, appeared to have overcome these obstacles.

Table 1
Prospective Elementary Teachers’ Responses to Relevant Survey Items at the Beginning and at the End of the Course (n=39)

	Mean (SD)		Result of comparison[b]		
Survey item[a]	Beginning	End	t(36-38)	p	Cohen’s d
1. Proofs are sequences of steps presented in two columns: one has the statements and the other has the reasons.	1.42 (.65)	2.44 (.85)	5.44	<.001	1.13
2. Proof is a topic that relates only to geometry	2.95 (.80)	3.74 (.45)	6.00	<.001	.81
3. Proof is an advanced topic that goes beyond the scope of the elementary school grades	2.22 (.75)	3.28 (.84)	7.22	<.001	.90

Notes. [a] Following Schoenfeld (1989), the Likert scale options for each item were “1: very true,” “2: sort of true,” “3: not very true,” and “4: not at all true.”
[b] The effect size d was large (>.80) for all three items.

To promote a pedagogical need among prospective teachers to learn about proof, we used two complementary approaches. The first approach utilized a key idea emerging from research on *mathematical knowledge for teaching* (Ball et al., 2008), namely, the importance of using pedagogical situations to contextualize prospective teachers' learning of mathematics in ways that have a bearing on the work of mathematics teaching (Heid & Wilson, 2015; Stylianides & Stylianides, 2014a; Wasserman et al., 2019). We operationalized this idea in a special kind of tasks that we call *pedagogy-related mathematics tasks* (Stylianides & Stylianides, 2014a) and that have two major features: a *mathematical* focus that relates to mathematical ideas that theory, research, or practice suggest are important for teachers to know; and a substantial *pedagogical context* that is an integral part of the task and essential for its solution. By interweaving mathematics and pedagogy, Pedagogy-Related Mathematics (PR-M) tasks not only can facilitate learning of mathematics that is useful in and for teaching (Ball & Bass, 2000), but also, and more relevantly to our purposes here, can help nurture a pedagogical need for that learning (Stylianides & Stylianides, 2014a).

As part of the instructional sequence for milestone 1b, we used the two PR-M tasks outlined in Fig. 4, which we asked prospective teachers to complete individually as part of a homework assignment before the next class session. To design the *pedagogical context* of these tasks, we used a classroom episode reported in Zack (1997). Vicki Zack, a teacher-researcher, described her Canadian fifth-graders' encounters with the notion of proof in the context of a version of the Squares Problem that was almost identical to the version we had used previously with the prospective teachers (Fig. 3) for promoting milestone 1a.[6] The episode presented Zack's elementary students engaging in mathematics as a sense-making activity, with proof playing a key role in their work: the students had a disagreement over a method for finding the total number of squares in a 60-by-60 square, and they sought to resolve

[6] The version of the Squares Problem used by Zack (1997) included an additional question about the number of different squares in a 10-by-10 square. Also, it did not ask students to explain or indicate their confidence for their answer for the number of different squares in a 60-by-60 square.

their disagreement by means of proof ("You have to prove us wrong") rather than by appeal to an authority. We expected that the fact that the words "prove" and "proof" were used naturally and meaningfully by the elementary students in the context of the same problem that the prospective teachers had worked on previously would reinforce for prospective teachers the relevance of proof to elementary school students' work, thereby helping promote both aspects of pedagogical need under milestone 1b.

Pedagogy-Related Mathematics (PR-M) task 1

The pedagogical context of PR-M task 1 derived from an episode reported in Zack (1997, pp. 294-295) that we presented to the prospective teachers as follows:

> The students in Vicki Zack's fifth-grade class are working in small groups on the "Squares Problem." **Will, Lew, and Gord (group of 3)** meet with **Ross and Ted (group of 2)**, and form a group of 5. Although their approaches to different parts of the "Squares Problem" have varied, the five peers have all been in agreement with the answers for the number of squares up to the 10 by 10 case. The answer for the number of squares in a 4 by 4 is 30, in a 5 by 5 is 55, and in a 10 by 10 is 385. [*Note: these answers are correct.*]
>
> During the discussion of the number of squares in a 60 by 60 square, Will, Lew, and Gord say that they have not yet completed getting an answer. **Ross and Ted, however, feel that they have the solution for the 60 by 60, which is to take the 385 (the answer for the 10 by 10 square) and multiply it by 6 to get the number of squares for a 60 by 60; the resulting answer is 2310.** [*Note: 10 of the 26 children in the class used this strategy.*] Will and Lew are very sure that Ross and Ted are wrong:
>
> Lew: I'll make you a bet.
> Will: I'll make you a bet.
> Lew: I'll bet you anything in the world.
> Ross: I'm not betting. **You have to prove us wrong.**

The mathematical focus of the task asked prospective teachers to respond to the following questions:

> Which group do you think is right? Why? What would you do to **prove the other group wrong**? Try to come up with a valid mathematical argument.

Pedagogy-Related Mathematics (PR-M) task 2

The pedagogical context of PR-M task 2 was the same as in PR-M task 1 above with the following addition: we presented to the prospective teachers three arguments that Gord, Lew, and Will developed (as reported in Zack, 1997) in order to prove Ross and Ted wrong. The mathematical focus of the task asked prospective teachers to judge the validity of each student argument and explain their thinking.

Fig 4 Outline of two "pedagogy-related mathematics tasks" (Stylianides & Stylianides, 2014a) used as part of the instructional sequence for promoting milestone 1b

The *mathematical focus* of PR-M task 1 in Fig. 4 was for prospective teachers to construct a proof to resolve students' disagreement, while the mathematical focus of PR-M task 2 was for them to evaluate arguments that elementary students constructed for that purpose. While the prospective elementary teachers were not yet introduced to criteria for proof, their engagement with these questions was intended to prime them for milestone 2 in the learning trajectory. Also, the consideration of the elementary students' debate in PR-M task 1 and their arguments in PR-M task 2 aimed to emphasize further the relevance of proof to school mathematics (aspect 1 of pedagogical need) and the ability of elementary students to engage meaningfully with this notion (aspect 2).

The second approach we used to promote a pedagogical need among prospective teachers complemented the approach we just described (based on PR-M tasks) and utilized the following key idea that emerged from research on teachers' professional learning: the use of narrative or visual cases (also referred to in the literature as vignettes, scenarios, stories, comic-style animations, etc.) of teaching practice to offer a concrete context for (prospective) teachers to project themselves into a certain pedagogical context and reflect on pedagogical or mathematical issues pertaining to that context (e.g., Arbaugh, Smith, Boyle, Stylianides, & Steele, 2018; Herbst & Chazan, 2011; Skilling & Stylianides, 2020). Although there are many published narrative or visual cases, none of these were suitable for our aim to create a pedagogical need for proof among prospective elementary teachers. Engaging elementary students in proving is uncommon in school mathematics practice internationally (Stylianides et al., 2017) and so relevant cases based on actual classroom data at the elementary school level are scarce; the few available cases are found mostly in research reports of teacher-researchers' practices (e.g., Ball & Bass, 2000) and are not packaged in a way that is readily usable in teacher training. Even fictional cases in the area of proof are hard to find for this particular school level; the only purposefully created narrative or visual cases related to proof we know of focused on the secondary school level (Arbaugh et al., 2018; Herbst & Chazan, 2011).

Despite these difficulties, we found what we were looking for in Zack's (1997) paper that we referred to earlier and that constituted the basis for the design of the two PR-M tasks. This was a brief conference paper, which, although not written as a narrative case, was essentially an authentic, first-person account of an elementary teacher who critically reflected on her proof-related instruction with the Squares Problem, having already endorsed and articulated in her paper the ideas that proof is relevant to elementary school mathematics (aspect 1 of pedagogical need) and that elementary students can engage meaningfully with proof (aspect 2). Another reason for the suitability of Zack's paper for our purposes was that, while Zack endorsed the two aforementioned ideas, she also discussed openly some mathematical and pedagogical challenges she faced when trying to engage her elementary students in proving; this honest, non-idealized reflection made it more likely that our prospective teachers would project themselves into Zack's position (cf. Skilling & Stylianides, 2020). Also, Zack's reflective account was consistent with the three considerations for communal criteria for proof, which we planned to guide our work in milestone 2 (Fig. 1). For example, regarding consideration 3, Zack expressed the contention "*that in order for an argument to be considered a proof, the students need not only* ***convince****, but also to* ***explain***" (p. 291; italics and emphasis in original).

Given the suitability of Zack's paper as a narrative case for our aim to promote milestone 1b among our prospective teachers, we asked them to read Zack's paper as part of the same homework assignment that included the two PR-M tasks. Also, we asked the prospective teachers to respond to a new CAP that prompted them to identify and rank order three ideas in Zack's paper that they considered important, and explain briefly why they considered them to be important. We intentionally did not channel prospective teachers' choices towards ideas relevant to milestone 1b; our experience using this CAP in previous research cycles suggested that the bulk of what prospective teachers would choose to comment on would still be in the area of proof. Also, we expected that their choices would constitute a good basis for the whole class discussion that the instructor would orchestrate in the next class session with the aim to "institutionalize" (Brousseau, 1981), that is, raise and give official status to, the two aspects of pedagogical need under milestone 1b.

In what follows, we discuss transcript excerpts from the whole class discussion of prospective teachers' selected ideas from Zack (1997) in one of the university classes (18 prospective teachers), which illustrate how the two aspects of pedagogical need can be institutionalized; the discussion in the other university class was similar. The instructor orchestrated the whole class discussion by implementing the four key practices we discussed in Section 3.2. In using practice 4 – revoicing or highlighting selected student contributions – the instructor kept in mind not only the two aspects under milestone 1b but also the three considerations for the proof criteria that would be introduced later under milestone 2.

The discussion started with prospective teachers sharing key ideas they identified in Zack (1997). Maria and Amanda were the first ones to make comments that directed the discussion to the notion of proof, which remained the focus of the discussion thereafter. None of the prospective teachers questioned Zack's pedagogical decision to engage her elementary students with proof. Rather, this decision seemed to be readily accepted by the prospective teachers who focused their comments on the learning affordances of elementary students' engagement with proof.

Maria: I also thought that one of the most important ideas was, like at the end, the teacher [Zack] went back to make sure that everyone understood, and that having them [the students] explain it [the process of finding the number of different squares] helped.

[…]

Amanda: I think it was important that she [Zack] didn't assign the task [the Squares Problem] with the intention of talking about proof, but the kids were curious and led the discussion there. She went [along] with it, and allowed them to deepen the ideas on their own.

Maria and Amanda's comments linked elementary students' engagement with proof with deepening students' understanding of mathematical ideas. Maria's comment also suggested a connection between students' understanding and them explaining their thinking. Amanda's comment suggested further that not only can proof be relevant to elementary students' mathematical work but also it can emerge naturally from that work, thereby reflecting an appreciation of both aspects of pedagogical need.

To draw attention to the explanatory function of proof, the instructor asked prospective teachers to comment on the elementary students' reactions when Zack showed to them the following formula for the sum of the first N square numbers: $N\cdot(N+1)\cdot(2N+1)/6$. Zack found this formula in Anderson (1996) but noted she did not know how the formula was derived. Sherrill's comment reinforced the explanatory function of proof as a "credible" means for establishing knowledge in the classroom as opposed to "given" knowledge by an authority (in this case, a journal article):

Sherrill: They [the students in Zack's class] basically like wanted him [Anderson] to explain *why*[7] it [the formula] worked. It looked like it would work, it looked like a good idea but so they didn't find it as *credible* as if she had explained why. This brought up the whole idea of proof, because they were just *given* the formula [from Anderson's article].

Prospective teachers' recognition that proof can be relevant to school mathematics (aspect 1 of pedagogical need) was further suggested by their subsequent discussion of the importance of teachers doing tasks like the Squares Problems

[7] *Italics* in transcript excerpts are ours to draw attention to specific segments.

themselves and being able to explain why things work when they implement these tasks with their students. Natalie's comment illustrates this point:

Natalie: One of the things that I said was important was *the teacher's role in explaining how it works*. And she [Zack] said that the way she did that was by talking about her own inquiries and her own steps through solving the problem. She didn't just say: "Okay I need to solve this Squares Problem, look in a book and find this formula." She went through and did it herself as well so she could see what the students were talking about.

The instructor then added to Natalie's comment that, if a teacher only knows one way to solve a problem, the teacher will not be well equipped to respond to different student arguments, such as the arguments that the prospective teachers had analyzed in PR-M task 2 (Fig. 4).

The discussion shifted next to another important idea in Zack's paper, namely, that an argument should convince not only oneself and a friend but also a skeptic. As we noted earlier, these three levels of conviction by Mason (1982) are relevant to the justification function of proof (consideration 3 under milestone 2).

Beth: I just thought that it was important because Will [a student in Zack's class] was determined to find a pattern, so that's the only way he wanted to look at it, and didn't really have an open mind to it. And I don't really get how he convinced his friend, he just said you know, the pattern's right, and he thought he was a genius. […]

Stylianides: And what's harder to do? To convince yourself, to convince a friend, or to convince a skeptic?

Beth: Convince a skeptic.

Knowing that these levels of conviction would be part of the conceptualization of proof the prospective teachers would be introduced to in milestone 2, Stylianides commented that the three levels of conviction were relevant to their teacher education class as much as they were relevant to Zack's elementary school class. After that, he asked directly whether it would be meaningful to talk about "proof" in the elementary grades, a question that related to both aspects of pedagogical need.

Joan responded in the affirmative and explained how during the previous activities for milestone 1a (outlined in Fig. 2) her thinking shifted from a restricted view of proof as an argument in geometry presented in two columns (cf. survey items 1 and 2 in Table 1) to a view of proof as a means for engaging children in sense-making across mathematical domains:

Joan: I had a pretty strict interpretation of proof before I came in here, which was the geometry proof [in high school] which I don't really remember very well, but I know that there's a proof using geometry; there's a column for the steps and a column on the other side for why you're doing that and the reasoning behind it. But so yeah, this is an interesting way, I guess, *to explain, the students or children themselves [in elementary school], why something works and it doesn't have to be specific to geometry or the proofs in geometry*.

Stylianides: [To the class:] Did you write something on this issue? [Pause. Some prospective teachers nod in agreement.] So what does Vicki Zack take proof to mean? What is proof for her? Amanda?

Amanda: Um, explain. She takes it to mean as long as you can *explain* why you're doing that and why it works. That's what it seems to be.

Building on Joan and Amanda's comments, Stylianides ratified the connection between the notions of "explanation" and "proof," and he offered a fuller account of Zack's (1997) view of proof to refer also to "conviction." He concluded the discussion by giving an official status to the idea that emerged from the previous contributions (relevant to milestone 1b) that proof can be a meaningful concept for elementary school mathematics and thus important for prospective elementary teachers to learn about, provided that proof is appropriately conceptualized. This comment led smoothly to the class's work on milestone 2.

4.3. Promoting milestone 2: an operationally functional conceptualization of proof

In considering learners' introduction to an operationally functional conceptualization of proof, one needs to recognize that it is unrealistic for instructors to expect learners to discover their own proof criteria (Hanna & Jahnke, 1996; Mariotti, 2006). Thus, instructors have a critical role to play in offering to members of their classroom communities (students and prospective teachers alike) access to appropriate criteria for proof. Yet it is not a matter of instructors simply giving to learners a set of proof criteria for learners to accept and apply. Rather, instructors should provide opportunities for the criteria to be socially negotiated and understood in the classroom community – both at the point of the introduction of the criteria (as in Campbell & King, 2020; Yee et al., 2018) and over time during their application in proving tasks (as in the two design experiments) – while ensuring that the criteria do not compromise important considerations like those under milestone 2 (Fig. 1).

Learners' readiness to make sense of or accept the instructor's proposed criteria is fundamental. Take, for example, a possible criterion for proof to be a convincing argument: a student who views empirical arguments as proofs would have different standards for conviction (for that student, empirical arguments *are* convincing) than another student who has been supported to overcome the epistemological obstacle of viewing empirical arguments as proofs (milestone 1a).

To introduce the participants in our studies to an operationally functional conceptualization of proof, we took into account the aforementioned ideas about the instructor's role. Also, in both studies, we presented the participants with a starting set of criteria, and we organized a discussion around the proposed criteria so as to socially negotiate the criteria and institutionalize them for further negotiation and use in future classwork.

Next we discuss how we introduced the participants in each study to communal criteria for proof. In doing so, we offer concrete images of what the social negotiation of communal criteria for proof can look like at the point of the introduction of these criteria, an issue that received little attention in prior relevant studies (Campbell & King, 2020; Stylianides, 2019; Stylianides & Stylianides, 2009a; Yee et al., 2018). Two major similarities between the starting sets of proof criteria used in each study were that the criteria (1) were consistent with the three considerations under milestone 2 and (2) were viewed as tentative and subject to further negotiation and discussion as each class continued to engage with proof over time. A major difference concerned how the proof criteria were devised in each study. The criteria for the secondary students were purely our own creation, in collaboration with the two secondary teachers, and so we had flexibility to devise criteria that would meet squarely the three considerations. We did not have the same flexibility in devising the starting set of criteria for the prospective teachers: a textbook we used occasionally in the course had a chapter on characteristics of a "good explanation" in mathematics (Beckmann, 2005) that, although not ideal for our purposes, we felt we could productively incorporate into our instructional design as we explain below.

4.3.1. The case of prospective teachers

The discussion of Zack (1997) that we reported previously showed that the prospective teachers were prepared to accept several criteria for proof, such as criteria relating to proof's explanatory and convincing functions (consideration 3). In preparation for a more focused discussion of proof criteria, we asked the prospective teachers to read Beckmann's (2005) discussion of characteristics of a "good explanation" in mathematics (first column in Table 2) and respond in writing to the following prompt: "Does Beckmann's list of characteristics of *good explanations* make sense to you? Is there anything you would like to add, delete, or change?"

Although Beckmann discussed criteria for explanations rather than proofs, we could still use her list and help the class view "proof" as a special type of a "good explanation" (in Beckmann's sense) applicable to tasks that required the solver to show the truth or falsity of a statement; this view of proof would be consistent with some of the literature that described proofs as a special kind of "explanatory arguments" (Balacheff, 1988; Knuth, 2002). The other major function of proof we decided to focus on – conviction – also featured prominently in Beckmann's characteristics 3b and 3c, thus addressing consideration 3 for proof criteria.

Table 2
Characteristics of a Good Explanation in Mathematics and the Criteria for Proof Used in the University Study

Beckmann's (2005) characteristics of a good explanation in mathematics	Characteristics of a good explanation as adapted by the prospective teachers for use in the teacher education class (These served also as *criteria for proof*)
1. The explanation is factually correct, or nearly so, with only minor flaws (for example, a minor mistake in a calculation).	1. The explanation is correct.
2. The explanation addresses the specific question or problem that was posed. It is focused, detailed, and precise. There are no irrelevant or distracting points.	2. The explanation addresses the specific question or problem that was posed. It is focused, detailed, and precise. There are no irrelevant or distracting points.
3. The explanation is clear, convincing, and logical. A clear and convincing explanation is characterized by the following:	3. The explanation is clear, convincing, and logical. A clear and convincing explanation is characterized by the following:
(a) The explanation could be used to teach another (college) student, possibly even one who is not in the class.	(a) The explanation uses language, representations, definitions that are understood by the people to whom the explanation is addressed.
(b) The explanation could be used to convince a skeptic.	(b) The explanation could be used to *convince a skeptic*.[a]
(c) The explanation does not require the reader to make a leap of faith.	(c) The explanation does not require the reader to make a leap of faith (e.g., "This is how it is" or "You need to believe me").
(d) Key points are emphasized.	(d) Key points are emphasized.
(e) If applicable, supporting pictures, diagrams, and equations are used appropriately and as needed.	(e) If applicable, supporting pictures, diagrams, and equations are used appropriately and as needed.
(f) The explanation is coherent.	(f) The explanation is coherent.
(g) Clear, complete sentences are used.	(g) Clear, complete sentences are used.
	(h) The explanation could be used by someone to solve a similar problem.

Note. [a] The prospective teachers wanted this phrase to be in italics for emphasis.

The other two considerations were also satisfied to a good extent in Beckmann's list. Regarding consideration 1 (honesty to the discipline of mathematics), Beckmann required good explanations to be factually correct (see characteristic 1, even though this characteristic leaves open the possibility for minor errors) and to address precisely, clearly, convincingly, and logically the question at hand (characteristics 2 and 3). Regarding consideration 2 (contextual appropriateness), Beckmann required good explanations to be formulated with other members of the classroom community or a similar kind of audience in mind (characteristic 3a). If we had authored the characteristics, we would have modified some of them, but overall we felt that Beckmann's list offered a good *starting* point for our discussion with the prospective teachers. Also, the fact that the list could be used as a pre-discussion reading offered the opportunity for prospective teachers to reflect individually on the criteria and come prepared for the discussion.

The second column in Table 2 presents the criteria for proof agreed upon in the class following the whole class discussion of Beckmann's list. As we can see in the table, several criteria were modified in that discussion. Below we present excerpts from the whole class discussion to illustrate how some of selected criteria were negotiated and accepted. During the discussion the instructor kept a public record of what was being agreed by annotating Beckmann's list. To identify the points of consensus and to steer the discussion towards changes we thought would strengthen the alignment of the accepted criteria with the three considerations, including our intended link between "proof" and Beckmann's notion of a "good explanation," the instructor orchestrated the whole class discussion by implementing the four practices we described in Section 3.2.

Characteristic 2 in Beckmann's (2005) list sparked considerable discussion.

Laura: On number 2, I think the statement "there are no irrelevant or distracting points" is subjective. Some kids might have to go through the irrelevant or distracting points to eliminate those as part of the thinking process. […]

Stylianides: This is an interesting point. Also in number 2 it says that the explanation should be "detailed" and then it says that there are "no irrelevant or distracting points." Do you see any conflict with these two? Lindsey?

Lindsey: Yeah, I also thought something similar to her [Laura], that they might have to go through… but I thought more, like, I guess I thought more of the teacher bringing up the distracting point. I guess I read it that way, like bringing up a point from a previous lesson that might need to be brought up to further explain and understand. […]

Natalie: Well, I read it as "detail" and "to the point" where it's necessary. If you think about the example that, if you're teaching a college-level course, you wouldn't need to explain how to do a subtraction problem. Like that would be detail, but it wouldn't be necessary, it would be irrelevant and distracting.

Stylianides: So what both of you [Lindsey and Natalie] are saying and what Laura said earlier, that it depends on the students, is perhaps one characteristic that is perhaps missing from this list: that explanations should be suitable for the *audience*, for the students who will be using them, and their *background knowledge*. What do you think about that? Someone noticed that? Britni?

Natalie's comment was reconciling Laura and Lindsey's earlier points and illustrates how the meaning of characteristic 2 was socially negotiated in the class. In his last comment, Stylianides implemented practice 4 so as to highlight some relevant bits in Lindsey and Natalie's contributions and direct the class's attention to issues of audience and contextual appropriateness (consideration 2). Britni's response to the instructor's questions led to revision of characteristic 3a:

Britni: I was going to say about… which one is it… [characteristic] 3a where it's saying that it should be irrefutably taught to someone else. Not all kids learn the same way or explain in the same way, so if you're talking about younger kids, like younger kids to explain this as opposed to an adult.

Stylianides: So then I will change that [characteristic 3a] and say that a good explanation should be appropriate for the background knowledge and understanding of the students. Okay? [Britni nods in agreement.] This issue about the detail and non-relevant and distracting points […] was similar to what Natalie said earlier: it should be detailed as long as it makes sense. Detailed enough for the person who is the audience. […]

The previous exchange is illustrative of the important role of the instructor in ensuring that the list of characteristics captured appropriately the three considerations while listening carefully to and building on prospective teachers' contributions. In addition to capturing better consideration 2, the revised characteristic 3a itemized aspects of an explanation (language, representations, definitions) that require attention when accounting for the audience of an explanation, thereby operationalizing the characteristic and strengthening its usability.

The prospective teachers made some additional comments on the other points under Beckmann's characteristic 3. For example, characteristic 3h was added following a suggestion from Laura and agreement by the rest of the class.

Laura: On number 3, I also added that the explanation that is given could be used to solve a similar problem where maybe just some of the numbers were changed. That could be another way of them, you know, offering *proof*.
Stylianides: What do other people think about that? So how would you write that, Laura?
Laura: That the explanation can be used to solve another variation of the problem, is what I wrote.
Stylianides: Do you think… should we add it to the list? [Prospective teachers nod in agreement and Stylianides adds it to the list.]

The previous excerpt illustrates the instructor's readiness to accept inclusion of criteria to the list that we did not consider necessary but were still viewed as important by the classroom community and did not compromise the three considerations. Laura's comment illustrates also that some prospective teachers were already thinking of the list of characteristics for good explanations as applying also for proofs. The instructor picked up on Laura's reference to "proof" and made explicit for all prospective teachers the connection between the notions of a "good explanation" and a "proof." He also clarified the way in which the list of characteristics would be used in the future work of the teacher education class.

Stylianides: Okay, so we can always revisit this list later in the course […] It would be good for us if we could try to follow the characteristics that appear on this list that are applicable in the problem that we're trying to solve or in homework assignments. […] Before we move on to something else, I would like to say something about the relation between a "good explanation" and a "proof," as we discussed in the context of the Vicki Zack paper last week. The same criteria would apply also for a good proof. The only difference in my mind between a good explanation and a good proof is that "proof" is a more restricted term that applies when we want to show whether a statement is true or false. So the explanation could be [a] broader [term] and include also other things that go beyond showing whether something is true or false. But I think that perhaps all of the characteristics that we have here would also be applicable to a good proof. What do you think? […]
Natalie: Well, I think that it would be just as necessary for a proof to be answering a specific question, logical explanation… which also applies to a proof.

Natalie's comment reinforced the idea proposed by the instructor that the same characteristics would apply for proofs. Once it was established that the revised list of Beckmann's characteristics of a good explanation would be used also as proof criteria in the teacher education class, the instructor made a note about that in the document he was annotating publicly so as to "institutionalize" (Brousseau, 1981) this piece of knowledge for use in the classroom community's future work.

Before we move on to the case of secondary students, we note that in the discussion of characteristics 3a and 3b, the prospective teachers made references to earlier parts of the intervention (the tasks in Fig. 2, the PR-M tasks in Fig. 4, and the discussion of Zack's paper) thereby illustrating their understanding or indicating their acceptance of those criteria. Due to space constraints, we do not elaborate on the discussion of those characteristics, but the reader can see similar issues reflected in the next section in the discussion that took place in the secondary classroom.

4.3.2 The case of secondary students

In accordance with the lesson plan that the researcher had agreed with the secondary teachers, the teachers capitalized on their students' emerging intellectual need to learn about proof (milestone 1a) to introduce them to a proposed set of criteria for proof. The criteria, which we present in Fig. 5 in the form of the actual PowerPoint slide that was used by the teachers, were developed collaboratively between the researcher and the teachers. The researcher brought into play his knowledge of the relevant literature (including the three considerations) and prior experience from the university study, while the two secondary teachers offered useful practical knowledge so that the criteria were contextually appropriate (consideration 2), notably understandable to and usable by their students. Contextual appropriateness was also reflected in criterion 3, which essentially referred to the community's "set of accepted statements" (Stylianides, 2007a). Regarding honesty to mathematics (consideration 1), an argument that met the standard of proof had to contain no errors (criterion 4), be clearly presented (criterion 5), and not require one to make a leap of faith (criterion 1). The latter requirement was meant to emphasize logical thinking as the basis of epistemic belief rather than appeal to an authority. Relatedly, an argument qualifying as a proof had to be convincing and explanatory (criteria 1 and 2, respectively) thus addressing both main functions of proof in consideration 3.

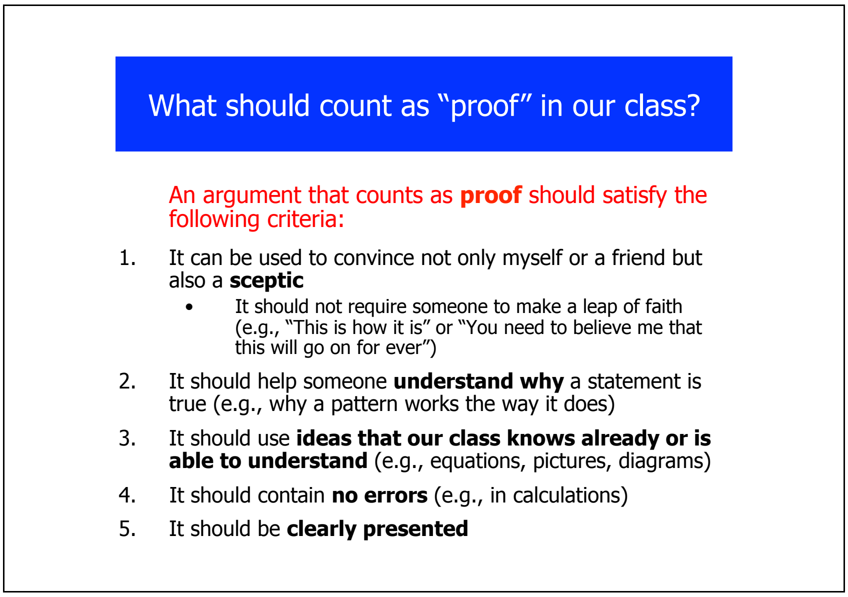

Fig 5 The criteria for proof that the teachers proposed to their students in the school study

In what follows we briefly illustrate the introduction of one of the secondary classes (30 students) to the proposed criteria for proof; a similar discussion took place in the other class. According to the plan, the classroom teacher, who we call Kathy (pseudonym), would solicit students' own views about criteria for proof before presenting the proposed criteria in Fig. 5 for the students to comment on and begin to develop a *shared* understanding of their meaning. The teacher would use the same four practices as the university instructor when orchestrating the whole class discussion, and she was prepared to accept modifications the students would suggest to the criteria that would not compromise the three considerations.

Kathy: […] We need to have some sort of idea then of what we need our proof to consist of. So let's have a think about the sorts of things that we need our proof to do. What for us is going to be enough to guarantee that we have proved it [a statement], guarantee that that the pattern will always work. What might our proof consist of?

Kathy's question solicited first a comment from Larry that related to proof's explanatory function (consideration 3):

Larry: It needs to show *why* there are actually 16 3-by-3 squares in a 5-by-5 grid. [Larry refers to question 2 in the Squares Problem. He probably meant to say "9" instead of "16," which was the result established in the class for this question.]

Kathy revoiced Larry's point (practice 4) given its relevance to criterion 2 (Fig. 5) and invited further contributions from the class:

Kathy: We have to show *why*. We can't just say, "this is the formula and it works." We have to show *why* it works. We have to show why there are so many 3-by-3 boxes in a 5-by-5, 6-by-6, or whatever… So we have to help people reading our proof *understand* it, it has to be understandable. What sorts of things might our proof have in it, then? Especially the proof for the Squares Problem, what might it look like when it gets on paper?
Robert: It might have a diagram in it.
Kathy: What else might have in it?
Larry: You would put in *how* you got to the equation yourself.
Kathy: Yes, you could have some neat working to show the steps you went through to get the final, whatever it may be, a formula or explanation.

Larry's new contribution seemed to relate to the presentation of a proof and the importance of proof making relevant processes explicit (consideration 1).

Notable in the discussion so far were students' references to earlier parts of the intervention. At this point, Kathy presented the PowerPoint slide in Fig. 5, indicating that the criteria were not set in stone and the class could modify them or revisit them in the future. She read the criteria and commented briefly on each of them, engaging students in discussion about what the criteria meant for them and how they thought the criteria related to their earlier mathematical work. Here is an excerpt from the discussion concerning criterion 1:

Kathy: What do we mean by "skeptic"?
Blaze: Someone who doesn't believe us.
Kathy: So we've got to convince even those skeptical persons. And most importantly, it shouldn't require someone to make "a leap of faith," to just trust us, like "this is how it is" or "you need to believe me that this will work for ever," which is what

we've done so far, isn't it? [She refers to the empirical arguments the class had produced earlier in the Squares Problem.] […]

Students: Yes. [Some students say "yes"; others nod in agreement.]

Kathy: We checked cases and said, "Oh, we've seen it working, therefore it will work for the rest." And we showed that perhaps this isn't the case like in [Task 2, Fig. 2].

In this excerpt Kathy presented criterion 1 as one that would disqualify empirical arguments from meeting the standard of proof by explaining to the students that arguments of this kind would fail to convince a skeptic. This comment addressed an important aspect of consideration 1 and offered to the class a sense of what would dissatisfy a skeptic in terms of accepting an argument as a proof. The reference to "conviction" also made a connection with consideration 3.

To further facilitate reflection on and understanding of the criteria, Kathy asked the students which criterion they thought was the most important. Sylvia nominated criterion 2, saying that an argument should help "rationalize it [a statement], understand why something is true." Sylvia's emphasis on the explanatory function of proof was consistent with Larry's earlier contribution. Recognizing its importance, Kathy revoiced Sylvia's contribution and made a connection between explanation and conviction (consideration 3): "If you can explain it, you are more likely to convince a skeptic." Overall, the students' contributions in this part of the discussion suggested that the criteria were contextually appropriate (consideration 2).

Finally, Kathy invited the class to say whether there were any further criteria they wanted to add to the list or any criteria they wanted to take out of the list. The class was content with the proposed criteria, so Kathy concluded the discussion by "institutionalizing" (Brousseau, 1981) the criteria for use in the future work of the class: she summarized again the criteria and said that an argument that fulfilled the criteria would "likely be a really good proof."

5. Concluding remarks

In recent decades, there has been notable progress in the field towards clarifying the meaning of proof for research purposes. We believe it is now time for the field to debate also, and become clearer about, the conceptualization(s) of proof that *students in school* and *prospective teachers in teacher education* can be introduced to so as to practically guide their engagement with proof. The lack of a clear stance on possible ways to introduce learners of mathematics to an operationally functional conceptualization of proof creates a significant obstacle (not the only one) to proof gaining the place that the field envisions for it in everyday classroom activity. Also, it perpetuates the imbalance in the current body of research that places more emphasis on documenting students and prospective teachers' difficulties with proof as compared to seeking ways to instructionally improve the learning experiences available to them to be able to productively engage with proof.

In this paper we took a step towards addressing this gap – which lies in the intersection of research, theory, and practice – by discussing and illustrating a learning trajectory and a respective instructional sequence that we used to help secondary students and prospective elementary teachers in two different countries to move along the milestones in the learning trajectory. We argued that, if students and prospective teachers' introduction to proof is consistent with one another (in terms of both the learning trajectory and the instructional sequence), like we proposed it to be, this can facilitate a coherent approach to the treatment of proof in school and teacher education settings.

A notable aspect of our proposed learning trajectory is that students and prospective teachers are expected to see an "intellectual need" (Harel, 2008) to learn about proof *prior* to their introduction to an operationally functional conceptualization of proof. To create such an "intellectual need," we targeted the removal of participants' epistemological obstacle of viewing proofs as empirical arguments, which prior research showed is a significant stumbling block to proof learning (e.g., Education Committee of the European Mathematical Society, 2011; Harel, 2008; Harel & Sowder, 2007). Once this epistemological obstacle is removed, we argued, participants are conceptually better prepared to make sense of and accept criteria for proof, including a notion of conviction linked to non-empirical arguments. Studies that did not focus on creating among their researcher participants an intellectual need for proof as a non-empirical argument observed resistance in participants' immediate work to use their proof criteria to evaluate empirical arguments as non-proofs (Campbell & King, 2020; Yee et al., 2018). The participants in our studies, on the other hand, showed in their immediate (oral) proof constructions (Stylianides, 2019) and in their longer-term (written) proof constructions and evaluations (Stylianides & Stylianides, 2009a) good awareness of the fact that empirical arguments are not proofs. Also, Brown (2014), in one of her studies where she followed an approach to promote an intellectual need among undergraduate students that was similar to ours, found that students' non-empirical proof conception was long-lasting.

Another notable aspect of our proposed learning trajectory is that prospective teachers need to see further a pedagogical need to learn about proof before their introduction to an operationally functional conceptualization of proof. The notion of *pedagogical need* is a new notion we introduced in this paper to describe the need for (prospective) teachers to learn about proof from a teacher's standpoint. Creating such a need is important in the context of the proposed learning trajectory because prospective teachers' epistemological readiness to learn about proof does not immediately translate into their readiness to learn about proof from a teacher's standpoint (Stylianides & Stylianides, 2014a; Wasserman et al., 2019).

While we made a case for the particular learning trajectory we presented in this paper and the significance of its various milestones, being guided by theoretical perspectives that we made explicit in our discussion and by a practical need to be selective and keep the scope of the learning trajectory relatively modest, we invite researchers to debate whether other milestones deserve also inclusion or priority over our selected milestones in the learning trajectory. With regard to the instructional sequence that we discussed in this paper, we present it as a possible way to promote the learning trajectory and to show that our theoretical proposition can be put into practice meaningfully rather than as a fully successful or the only way to achieve the goal of introducing learners to an operationally functional conceptualization of proof. Indeed, future research can help improve the instructional sequence or provide alternative pathways to promoting its various milestones.

Regarding milestone 1a, Brown (2014) already suggested an alternative pathway, called the "Cultural, Non-Experiential Pathway," for how to help students develop skepticism towards empirical validations through the power of an authority. She described this pathway as follows:

> The *Cultural Non-Experiential Pathway* is a didactical pathway that arises when an authority within the collective dismisses a basis of belief but there is no experiential history that the collective has shared, which serves as grounds for this action. In this case, a way of reasoning is established as an expected practice through the power of an authority. The authority's warrant for introducing the practice is that it is a practice of the larger community—it is a part of their culture—and, therefore, must be adopted by those who wish to legitimately participate. (Brown, 2014, p. 327, italics in original)

We have two reservations about this pathway. First, if students do not develop an intellectual need that would motivate them to change their views about empirical validations, they may simply be showing obedience to the authority with the latter running the risk of becoming a habitual way of reasoning, that is, a new obstacle to learning. Second, in the early cycles of our university-based design experiment we trialed a version of the instructional sequence to promote milestone 1a that relied on the power of an authority and obtained poor results: the instructor emphasized repeatedly to students that empirical validations do not meet the standard of proof within the broader mathematical culture but, although the students did not challenge the instructor and went along with his lead to seek non-empirical arguments, the students showed resistance to change their empirical conceptions. Despite our two reservations, however, we believe that this pathway as well as others merit further exploration, discussion, and debate.

Regarding milestone 2, Campbell and King (2020) and Yee et al. (2018) discussed an alternative pathway for how to engage learners in developing and socially agreeing on proof criteria. In both of these studies, learners (secondary students in Campbell & King, 2020; university students in Yee et al., 2018) had an opportunity to openly share and discuss their views of what should count as a proof in their class before the instructor played an active role selecting from, or helping the class to narrow down, the extended list of possible proof criteria proposed by the classroom participants. We believe this way of engaging learners in developing and socially agreeing on proof criteria could offer a viable alternative to our way of promoting milestone 2 so long as the communal criteria for proof ultimately accepted by the class satisfied some key considerations like those we discussed earlier in the paper (i.e., honesty to the discipline of mathematics, contextual appropriateness, and support of mathematical sense-making through justification/conviction and explanation).

Our proposed learning trajectory and respective instructional sequence may also be used with elementary (not only secondary) students and with prospective secondary (not only elementary) mathematics teachers, though the instructional sequence would likely need to be tailored to each of these groups. Regarding elementary students, the tasks that we used to generate an intellectual need for proof in the school study might need some modification, especially Tasks 2 and 3 in Fig. 2 (for the actual tasks, see Stylianides & Stylianides, 2009b). The Squares Problem (Fig. 3) was used by Zack (1997) with elementary students, so it has good potential to be used without adaptation with students of a similar age. The phrasing of the proof criteria we used with secondary students (Fig. 5) might also need some modification so that they are understandable to younger students as per our "contextual appropriateness" consideration.

Regarding prospective secondary mathematics teachers, they might need less support than prospective elementary teachers to see a pedagogical need to learn about proof, not least because proof tends to have a more visible place in the secondary than in the elementary school curriculum. Indeed, teachers tend to see more the relevance of content that they will teach and, even more so, content that their students will be assessed on (e.g., Birenbaum, 2014). As we explained previously, however, the notion of "pedagogical need" in the learning trajectory should be considered both in relation to the conceptualization of proof to which prospective teachers will be introduced next in the instructional sequence and in relation to prospective teachers' current conceptions of proof.

Finally, we draw attention to the dual role of teacher educators to introduce prospective teachers to the notion of proof and to prepare them as future teachers of mathematics who, in turn, will introduce their own students to proof. Our proposed learning trajectory and respective instructional sequence can be used in the service of this dual role. First, teacher educators can implement the instructional sequence (or an adapted version thereof) to help prospective teachers progress along the learning trajectory in a way similar to our university study. Afterwards, teacher educators can discuss with their prospective teachers how prospective teachers can help their future students to progress along the same learning trajectory in a way similar to our school study. Towards this end, teacher educators can use as a “narrative case” a paper like the one by Stylianides (2009a), which was published in a practitioner journal and described the implementation of the instructional sequence with secondary students, in order to discuss with prospective teachers pedagogical issues pertaining to the instructional sequence and reflect on the teacher’s role.

References

Arbaugh, F, Smith, M. P., Boyle, J., Stylianides, G. J., Steele, M. (2018). *We reason & we prove for ALL mathematics: Building students' critical thinking. Grades 6-12*. Thousand Oaks, CA: Corwin.

Anderson, J. (1996). The place of proof in school mathematics. *Mathematics Teaching, 155*, 33-39.

Bailey, D. H., & Borwein, J. M. (2005). Experimental mathematics: Examples, methods, and implications. *Notices of the American Mathematical Society, 52*(5), 502–514.

Balacheff, N. (1988). 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.

Ball, D. L., & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In J. Boaler (Ed.), *Multiple* *perspectives on mathematics teaching and learning* (pp. 83-104). Westport, CT: Ablex Publishing.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389–407.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences, 13*(1), 1-14.

Beckmann, S. (2005). *Mathematics for elementary teachers.* Boston: Addison Wesley.

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.

Birenbaum, M. (2014). Conceptualizing assessment culture in school. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing assessment for quality learning* (pp. 285-302). Springer.

Brousseau, G. (1981). Problèmes de didactique des décimaux. *Recherches en Didactique des Mathématiques, 2,* 3–128.

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

Bruner, J. (1960). *The process of education.* Cambridge, MA: Harvard University Press.

Buchbinder, O., & McCrone, S. (2020). Preservice teachers learning to teach proof through classroom implementation: Successes and challenges. *The Journal of Mathematical Behavior, 58*, doi: 10.1016/j.jmathb.2020.100779

Campbell, T. G., & King, S. (2020). Eighth-grade students' use of communal criteria for collaborative proving. *Investigations in Mathematics Learning, 12*(2), 124-141.

Clements, D. H., Sarama, J., Baroody, A. J., & Joswick, C. (2020). Efficacy of a learning trajectory approach compared to a teach-to-target approach for addition and subtraction. *ZDM Mathematics Education, 52*, 637-648.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13.

Cohen, D., Raudenbush, S., & Ball, D. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis, 25*(2), 119–142.

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

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

Education Committee of the European Mathematics Society. (2011). Do theorems admit exceptions? Solid findings in mathematics education on empirical proof schemes. *EMS Newsletter, 82,* 50-53.

Gal, H. (2019). When the use of cognitive conflict is ineffective – problematic learning situations in geometry. *Educational Studies in Mathematics, 102,* 239-256.

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

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, Netherlands: Kluwer Academic Publishers.

Harel, G. (2008). A DNR perspective on mathematics curriculum and instruction. Part II: with reference to teacher's knowledge base. *ZDM Mathematics Education, 40*, 893-907.

Harel, G. (2013). Intellectual need. In K. R. Leatham (Ed.), *Vital directions for mathematics education research* (pp. 119-151). New York: Springer.

Harel, G., & Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 805–842). Greenwich, CT: Information Age.

Heid, M. K., & Wilson, P. S. (with Blume, G. W.) (Eds.). (2015). *Mathematical understanding for secondary teaching: A framework and classroom-based situations.* Charlotte, NC: Information Age Publishing & National Council of Teachers of Mathematics.

Herbst, P. G. (2002). 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., & Chazan, D. (2011). On creating and using representations of mathematics teaching in research and teacher development. *ZDM Mathematics Education, 43*, 1-5.

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.

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

Küchemann, D., & Hoyles, C. (2001-03). *Longitudinal Proof Project* (Technical reports for Year 8-10 surveys). London: Institute of Education.

Leikin, R., & Dinur, S. (2003). Patterns of flexibility: Teachers' behavior in mathematical discussion. *Proceedings of the 3rd Conference of the European Society for Research in Mathematics Education.* Bellaria, Italy. Retrieved from https://www.mathematik.uni-dortmund.de/~erme/CERME3/Groups/TG11/TG11_Leikin_cerme3.pdf

Lockwood, E., Caughman, J. S., & Weber, K. (2020). An essay on proof, conviction, and explanation: multiple representation systems in combinatorics. *Educational Studies in Mathematics, 103,* 173-189.

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.

Mariotti, M. A., Durand-Guerrier, V., & Stylianides, G. J. (2018). Argumentation and proof. In T. Dreyfus, M. Artigue, D. Potari, S. Prediger, & K. Ruthven (Eds.), *Developing research in mathematics education - twenty years of communication, cooperation and collaboration in Europe* (pp. 75-89). Oxon, UK: Routledge.

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

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

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

Norwegian Directorate for Education and Training. (2020). *Curriculum for Mathematics, Year 1–10: Core Elements.* Retrieved from https://www.udir.no/lk20/mat01-05/om-faget/kjerneelementer?lang=eng

Piaget, J. (1985). *The equilibrium of cognitive structures.* Chicago, IL: University of Chicago Press (Original work published 1975).

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.

Schoenfeld, A. H. (1989). Explorations of students' mathematical beliefs and behavior. *Journal for Research in Mathematics Education, 20*(4), 338-355.

Schoenfeld, A. H. (1994a). Reflections on doing and teaching mathematics. In A. H. Schoenfeld (Ed.), *Mathematical Thinking and Problem Solving* (pp. 53-70). Lawrence Erlbaum Associates.

Schoenfeld, A. H. (1994b). What do we know about mathematics curricula? *The Journal of Mathematical Behavior, 13*(1), 55-80.

Schwab, J. J. (1978). Education and the structure of the disciplines. In J. Westbury & N. J. Wilkof (Eds.), *Science, curriculum, and liberal education: Selected essays* (pp. 229-272). Chicago: University of Chicago Press.

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 Mathematics Education, 46,* 767–780.

Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education, 26,* 114-145.

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.

Skilling, K, & Stylianides, G. J. (2020). Using vignettes in educational research: a framework for vignette construction. *International Journal of Research & Method in Education, 43*(5), 541-556.

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

Stylianides, A. J. (2007b). 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. (2016). *Proving in the elementary mathematics classroom*. Oxford, UK: Oxford University Press.

Stylianides, A. J. (2019). Secondary students' proof constructions in mathematics: the role of written vs. oral mode of argument representation. *Review of Education, 7*(1), 156-182.

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

Stylianides, A. J., & Stylianides, G. J. (2014a). Viewing "mathematics for teaching" as a form of applied mathematics: Implications for the mathematical preparation of teachers. *Notices of the American Mathematical Society, 61*(3), 266–276.

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

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

Stylianides, G. J., Sandefur, J., & Watson, A. (2016). Conditions for proving by mathematical induction to be explanatory. *The Journal of Mathematical Behavior*, *43*, 20–34.

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,* 103-133.

Stylianides, G. J., & Stylianides, A. J. (2009b). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education, 40*(3), 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), 374-415.

Stylianides, G. J., & Stylianides, A. J. (2017). Research-based interventions in the area of proof: The past, the present, and the future. *Educational Studies in Mathematics, 96*(2), 119-127.

Stylianides, G. J., Stylianides, A. J., & Weber, K. (2017). Research on the teaching and learning of proof: Taking stock and moving forward. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 237-266). Reston, VA: National Council of Teachers of Mathematics.

Wasserman, N. H., Weber, K., Fukawa-Connelly, T., & McGuffey, W. (2019). Designing advanced mathematics courses to influence secondary teaching:

fostering mathematics teachers' "attention to scope." *Journal of Mathematics Teacher Education, 22*(4), 379-406.

Watson, A., & Mason, J. (2005). *Mathematics as a constructive activity: Learners generating examples.* Mahwah, NJ: Erlbaum.

Weber, K. (2010). Proofs that develop insight: Proofs that reconceive mathematical domains and proofs that introduce new methods. *For the Learning of Mathematics*, *30*(1), 32–37.

Weber, K. (2014). Proof as a cluster concept. In C. Nicol, S. Oesterle, P. Liljedahl, & D. Allan (Eds.). *Proceedings of the Joint Meeting of PME 38 and PME-NA 36* (Vol. 5, pp. 353-360). Vancouver, Canada: PME.

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.

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

Yee, S. P., Boyle, J. D., Ko, Y. Y., & Bleiler-Baxter, S. K. (2018). Effects of constructing, critiquing, and revising arguments within university classrooms. *Journal of Mathematical Behavior, 49,* 145-162.

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.

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

Zaslavsky, O., Knuth, E., & Ellis, A. (Eds.). (2019). The roles and uses of examples in conjecturing and proving. *The Journal of Mathematical Behavior, 53*, 242-347.

Zazkis, R., & Chernoff, E. J. (2008). What makes a counterexample exemplary? *Educational Studies in Mathematics, 68,* 195-208.