# Towards a comprehensive knowledge package for teaching proof: A focus on the misconception that empirical arguments are proofs

**Author:**
Andreas J. Stylianides[1]

**Affiliation:**
[1]Faculty of Education, University of Cambridge, United Kingdom

**Correspondence to:**
Andreas Stylianides

**Email:**
as899@cam.ac.uk

**Postal address:**
University of Cambridge, Faculty of Education, 184 Hills Road, Cambridge CB2 8PQ, United Kingdom

The concept of *proof* is central to meaningful learning of mathematics, but is hard for students to learn. A serious misconception dominant amongst students at all levels of schooling is that empirical arguments are proofs. An important question, then, is the following: What knowledge might enable teachers to help students overcome this misconception? Earlier research led to construction of a significant but rather incomplete 'knowledge package' for teaching in this area. Major elements of this knowledge package are summarised and its further development is contributed to by discussing a research-based instructional intervention found to be effective in helping secondary students begin to overcome the misconception that empirical arguments are proofs. Implications for mathematics teacher education are considered.

## Introduction

A large body of research shows that students of all levels of schooling (including high-attaining secondary students) 'prove' mathematical generalisations by using empirical arguments (e.g. Chazan, 1993; Coe & Ruthven, 1994; Healy & Hoyles, 2000; Senk, 1985; Sowder & Harel, 1998), whereby 'empirical arguments' are those that purport to establish the truth of a generalisation by validating it only in a proper subset of all possible cases in its domain. Yet, the fact that a generalisation is found to be true in some cases does not guarantee − and thus does not prove − that it is true for all possible cases. This is a fundamental difference between an empirical argument and the notion of proof in mathematics (Stylianides, 2007b). However, many students have difficulty understanding this distinction.

In addition to establishing conclusively and convincingly the truth of a generalisation, a proof can also help someone understand *why* the generalisation is true. A proof's potential to promote conviction (justification) and understanding (explanation) in part accounts for it being considered central to meaningful learning in mathematics (Ball & Bass, 2000, 2003; Hanna, 2000; Harel & Sowder, 2007; Stylianides, 2008; Stylianides & Stylianides, 2008). According to Harel and Sowder:

> [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.

> (Harel & Sowder, 2007, pp. 808–809)

Unless students realise the limitations of empirical arguments as methods for validating generalisations, they are unlikely to appreciate the importance of proof in mathematics (Stylianides & Stylianides, 2009). In order to achieve this learning objective, however, teachers must have good knowledge in the area of proof, for the quality of learning opportunities that students receive in classrooms depends on the quality of their teachers' knowledge (Ball, Thames & Phelps, 2008; Goulding, Rowland & Barber, 2002; Ma, 1999). What knowledge, then, might enable teachers to help their students begin to overcome the misconception that empirical arguments are proofs?

## Knowledge for teaching proof

Following Shulman's (1986, 1987) influential work on the nature of teachers' knowledge for teaching, a significant body of mathematics education research has begun to unravel the kinds of knowledge that teachers need for effective mathematics teaching. Shulman's work has stimulated at least two major and related research strands in the mathematics education literature on teacher knowledge.

The first research strand has examined the mathematical demands that mathematics teaching places on teachers' knowledge (e.g. Ball & Bass, 2000, 2003; Ball et al., 2008; Davis & Simmt, 2006; Ma, 1999), and has helped identify important mathematical ideas essential for teachers to know in

order to support student learning of particular mathematical topics or concepts, such as proof (for a study that focused on proof, see Stylianides & Ball, 2008). This strand informs the field's understanding of teachers' mathematical (subject matter) knowledge about proof, one of three broad kinds of teacher knowledge for teaching proof considered here.

The second research strand stimulated by Shulman's work has examined students' common ways of thinking about particular mathematical ideas, helping identify important student conceptions (including misconceptions) that are essential for teachers to know in order to help their students develop their current conceptions. The extended mathematics education research on students' conceptions of proof (e.g. Chazan, 1993; Coe & Ruthven, 1994; Healy & Hoyles, 2000; Knuth, Choppin, Slaughter & Sutherland, 2002; Senk, 1985; Sowder & Harel, 1998; Stylianides & Al-Murani, 2010) informs understanding of teachers' knowledge about students' conceptions of proof, the second broad kind of teachers' knowledge for teaching proof considered here. This kind of knowledge can be regarded as an example of an aspect of Shulman's notion of pedagogical content knowledge, namely, knowledge 'of what makes the learning of specific topics [in this case proof] easy or difficult' for students (Shulman, 1986, p. 9).

The third broad category of teacher knowledge considered in this article, pedagogical knowledge for teaching proof, has received significantly less research attention than the other two, even though also related to Shulman's work. Prior research has offered few insights into pedagogical practices that are important for teachers to know and implement in their classrooms to help students develop conceptions of proof that better approximate conventional understanding. Such fine-grained, domain-specific pedagogical knowledge in the area of proof can be regarded as an example of another aspect of Shulman's notion of pedagogical content knowledge, namely knowledge of 'the ways of representing and formulating the subject matter that make it comprehensible to others' (Shulman, 1986, p. 9).

Effective teaching of proof requires at least these three broad kinds of knowledge, which are interrelated. For example, good knowledge about students' conceptions of proof is not possible without a robust understanding of the mathematical ideas that underpin these conceptions. Similarly, a teacher's ability to implement pedagogical practices in the classroom to help students develop their conceptions of proof requires a robust understanding of what the students' conceptions are and other forms (more advanced) that these can take.

What follows is a summary of what the field of mathematics education currently considers important for teachers to know in relation to the three broad kinds of knowledge, focusing on the common student misconception that empirical arguments are proofs. This is used to discuss how the field might develop a more comprehensive knowledge package for teaching in this area. I use the term *knowledge package* to describe a cluster of related kinds of knowledge (about mathematics, students and pedagogy) that are important for teachers to have in order to teach effectively a particular idea in their classrooms.[1]

## Mathematical knowledge about proof

In our review of prior literature on teachers' mathematical knowledge about proof for teaching (Stylianides & Ball, 2008), we identified understanding the distinction between empirical arguments and proofs to be an important element of this knowledge. Unless teachers at all levels of schooling develop a good understanding of this distinction, it is unlikely that large numbers of students will overcome their misconception that empirical arguments are proofs. According to Martin and Harel:

> [i]f [*elementary*] teachers lead their students to believe that a few well-chosen examples constitute a proof, it is natural to expect that the idea of proof in high school geometry and other courses will be difficult for the students.
>
> (Martin & Harel, 1989, pp. 41–42)

Also, elementary teaching practices that promote or tolerate a conception of proof as an empirical argument instill mental habits in students that significantly deviate from conventional mathematical understanding in the field. Dewey (1903, p. 217) cautioned educators against such practices when he said that whatever the preliminary approach to learning, it should not inculcate 'mental habits and preconceptions which have later on to be bodily displaced or rooted up in order to secure proper comprehension of the subject'.

## Knowledge about students' conceptions of proof

A significant body of research has investigated students' conceptions of proof and developed various taxonomies of these (e.g. Balacheff, 1988; Sowder & Harel, 1998); these in turn inform our understanding of what is important for teachers to know. Good knowledge about students' conceptions of proof can help teachers evaluate their students' current understandings about proof, and thus prepare teachers to design instruction that aims to help students develop conceptions of proof that better approximate conventional understanding.

The following is a taxonomy (hierarchy) of students' conceptions of proof, presented in increasing level of mathematical sophistication (for elaboration, see Stylianides & Stylianides, 2009):

- **Naïve empirical conception:** Validating a mathematical generalisation (pattern, conjecture, etc.) by checking a proper subset of all possible cases in its domain, selected primarily on the basis of convenience.

1.Ma (1999) used the term *knowledge package* to describe networks of relationships amongst different concepts that a teacher must understand in order to make proper decisions about which concepts are required for the learning of other concepts, which concepts can be learned simultaneously, etcetera. Thus, Ma used knowledge packages to describe concept maps or organising structures within teachers' mathematical knowledge for teaching. My use of the term is broader than that: it can include, but is not limited to, concept maps that show the relationships between proof and other closely connected notions such as pattern, conjecture, and argument.

- **Crucial experiment conception:** Validating a mathematical generalisation by checking a proper subset of all the possible cases in its domain; these cases are not selected on the basis of convenience but rather based on some kind of rationale, notably the search of possible counter-examples to the generalisation.
- **Nonempirical conception:** Recognising empirical arguments as insecure methods for validating mathematical generalisations; being able to understand and write proofs.

Naïve empiricism and crucial experiment are two kinds of empirical arguments (Balacheff, 1988), the former implying no awareness of the possibility that the examined generalisation could be false. Accordingly, although both represent an empirical approach to validating mathematical generalisations, the first is less advanced than the second. The nonempirical conception includes several stages, the most advanced being students' ability to understand and write proofs. This article focuses on what might be considered to be the first stage of a nonempirical conception – recognising the limitations of empirical arguments as insecure methods of validation, and seeing the need to learn about more secure methods (i.e. proofs).

## Pedagogical knowledge for teaching proof

Prior research has developed a useful basis about general pedagogical practices for engaging students in mathematical reasoning, argumentation, and proof (e.g. Ball & Bass, 2000, 2003; Stylianides, 2007a, 2007b; Stylianides & Ball, 2008; Yackel & Cobb, 1996; Zack, 1997). However, this body of research needs to be developed further before it can inform design of effective instruction to help students develop more accurate conceptions of proof, including the nonempirical conception.

For example, in our earlier work (Stylianides & Ball, 2008), we elaborated on the importance of teachers using a variety of proving tasks that can offer students learning opportunities to develop an understanding of different proving strategies and reasoning skills. Although implementation of particular proving tasks in the classroom can support generation of different kinds of arguments, thus offering an opportunity to teachers and students to discuss the differences between empirical arguments and proofs, it is unclear how this discussion can be organised to help students *overcome* their deeply rooted misconception that empirical arguments are proofs. Indeed, prior research and practice have shown that addressing this misconception is a stubborn problem in mathematics education, even at university level (Goulding & Suggate, 2001; Martin & Harel, 1989). This problem cannot be addressed without a carefully designed instructional intervention.

By *instructional intervention* is meant a purposeful and cohesive collection of activities (and respective implementation strategies) for achieving particular learning outcomes. Ability to successfully implement an instructional intervention in the classroom to help students begin to overcome the misconception that empirical arguments are

proofs would be an essential element of teachers' pedagogical knowledge about proof and, by implication, an important complement to the existing knowledge package for teaching proof.

What could such an instructional intervention look like, and what demands would successful implementation thereof place on teachers' knowledge? A research-based instructional intervention found to be effective in helping students begin to overcome the misconception that empirical arguments are proofs is discussed in the next section. This will exemplify also the point that successful implementation of the instructional intervention requires good knowledge about mathematics and students – emphasising the inextricable relationships amongst the different kinds of knowledge comprising the knowledge package for teaching proof.

# An instructional intervention
## Background

In a four-year design experiment in an undergraduate mathematics course in the United States of America (for discussion on design experiment methodology, see e.g. Schoenfeld, 2006), Gabriel Stylianides and I developed an instructional intervention. We showed this to be effective in helping undergraduate students begin to understand the limitations of empirical arguments and to see an 'intellectual need' (Harel, 1998) for learning about more secure methods of validation (i.e. proofs) (Stylianides & Stylianides, 2009). Presented at a later stage in this article is how a secondary mathematics teacher implemented a modified version of the original instructional intervention in a Year 10 class in England (14–15-year-old students) to achieve the same learning goals.

Implementation of the intervention by the secondary teacher was part of a school-based project organised as a design experiment, and following up on the aforementioned university-based design experiment. I conducted the school-based project in collaboration with two secondary mathematics teachers in two classes in a State school in England over a period of two years (when the students were in Years 10 and 11). The project aimed to generate theoretical and practical knowledge about how secondary teachers can better support students' learning of proof, and focused on high-attaining students. Given the findings of prior research that even high-attaining secondary students in England possessed weak understanding of proof (Coe & Ruthven, 1994; Healy & Hoyles, 2000; Küchemann & Hoyles, 2001–2003), and limited research knowledge on how to improve students' understanding of proof, I considered it strategic to focus on a student population that would offer better chances for success (high-attaining students), as a first step in a long-term research programme.

The theoretical framework underpinning the design and implementation of the focal instructional intervention is elaborated on in Stylianides and Stylianides (2009, pp.

316–324). The implementation of the intervention by the secondary teacher lasted approximately 60 minutes, and extended over two consecutive lessons. The teacher followed detailed lesson plans that I prepared and discussed with her prior to the lessons; our discussion led to minor modifications of the lesson plans, primarily in order to accommodate her timetable constraints (Stylianides, 2009a).[2]

Whilst reading the description of implementation of the intervention, I invite the readers to consider this question: What knowledge was it important for the teacher to have in order to successfully implement the intervention in her class? I will return to this question later in the article.

## Overview of activities in the intervention

The intervention included three activities: the Squares Problem, the Circle and Spots Problem, and the Monstrous Counter-example Illustration (refer to Figures appearing respectively throughout the article). The teacher used these three activities to facilitate students' progression along a 'learning trajectory' (Simon, 1995; see also Clements & Sarama, 2004) from a naïve empirical conception, to a crucial experiment conception, to a nonempirical conception. Each of these three stages in the learning trajectory corresponds to a conception in the taxonomy discussed earlier.

### Activity 1: The Squares Problem

The hardest part of the Squares Problem (Figure 1) was part three, which asked students to find and explain their answer for the number of different 3-by-3 squares in a case that was difficult for them to check practically ($n = 60$). The teacher made sure that the students understood what the problem said, and then asked them to work on it in small groups.

In small groups the students identified the pattern that the number of different 3-by-3 squares in an $n$-by-$n$ square was given by the formula $(n - 2)^2$. They verified the pattern for $n = 4$ and $n = 5$, and some of them for $n = 6$ as well. Based on this confirming evidence, they concluded that the pattern would hold for all values of $n$, including $n = 60$. Thus the students validated the pattern empirically on the basis of naïve empiricism. The whole-group discussion that followed further illustrated dominance of the naïve empirical conception in the class, with all small groups answering the three parts of the problem using the formula $(n - 2)^2$.

After some discussion about the meaning of the formula, the teacher (following the lesson plan) asked the students each to write down their thoughts about whether and why they could be sure that applying this formula would give the correct answer for part three of the problem. Some illustrative responses from students in one of the small groups were as follows:

Because we have found a formula and tried it against smaller squares so we can make sure that the formula is right [Bob].

I am sure that this solution works because it worked for every one we did [Calvin].

I am sure that the answer is correct because it has been proved for a number of smaller grids [Dan].

(Stylianides, 2010a, pp. 10–11)

These comments were representative of those of the rest of the class. The students were convinced of the truth of the pattern on the basis of naïve empiricism; the pattern worked for the first few cases and so they felt it would also work for $n = 60$.

After the students' individual reflections, the teacher proceeded with the next item on the lesson plan, which was to summarise the students' written responses. The teacher's summary was based on a quick inspection of students' written responses as she circulated during their individual reflection time. Accuracy of the teacher's summary was confirmed by a more careful analysis of students' responses by both her and me at the end of the lesson:

*'I get a feeling that most of you have said, "Well, I think we have sort of answered this question that $58^2$ is the right answer: we have found a pattern by checking smaller grid sizes and then we have used that pattern, assuming that it would continue all the way up to 60-by-60." That's the stage where we are right now: we've seen a pattern working, somebody said they tried the 6-by-6 and it worked for that too, and so we continued our pattern up to the $58^2$.'*

(Stylianides, 2010a, p. 11)

The state of student conceptions of proof in the class, as described in the aforementioned summary (where naïve empiricism predominated), had been anticipated in the lesson plan and was consistent with findings of prior research on the issue.

After the teacher's summary, the class moved on to the next activity. According to the lesson plan, the issue about the correctness of the pattern in the Squares Problem would
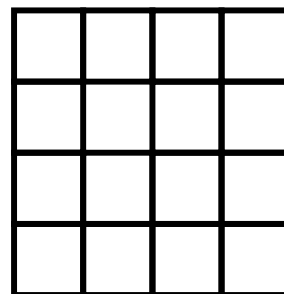


1. How many different 3-by-3 squares are there in the 4-by-4 square above?
2. How many different 3-by-3 squares are there in a 5-by-5 square?
3. How many different 3-by-3 squares are there in a 60-by-60 square? Are you sure that your answer is correct? Why?

*Source*: Adapted from 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

**FIGURE 1:** The Squares Problem.

---

2.The description of the implementation is a shortened and slightly adapted version of the description that was previously published in a teachers' journal (Stylianides, 2009a), and is used here with permission from both editors. Its use here is different and serves a new purpose, namely to explore issues about teachers' knowledge for teaching proof. Readers can also find the PowerPoint slides and lesson plans that the secondary teacher used in implementing the intervention in Stylianides (2009a).

remain tentatively unresolved. The class would revisit and resolve the issue after the students had been assisted to realise the limitations of empirical arguments (both naïve empiricism and crucial experiment). Our intention was for the students to realise the limitations of empirical arguments on their own, after experiencing a series of 'cognitive conflicts' (see Stylianides & Stylianides, 2009) and reflecting on situations in the next two activities, where empirical methods of validation were inadequate (for the readers' information, I note that the $[n - 2]^2$ pattern in the Squares Problem was actually correct).

**Activity 2: The Circle and Spots Problem**
The teacher introduced the Circle and Spots Problem (Figure 2), and helped the students understand what it said. She discussed with them the meaning of the terms *maximum* and *non-overlapping regions*, and clarified for them that *spots* meant points and that the phrase *around the circle* referred to the circle's circumference. She also mentioned that the points on the circumference did not have to be equidistant. Then the teacher asked the students to work on the problem in their small groups.

In our planning it had been anticipated that, similar to what students had done in the Squares Problem, they would check simpler cases, identify a pattern, trust the pattern based on naïve empiricism, and apply it to offer a definite answer for $n = 15$ (where $n$ stands for the number of points). However, there is a key difference between the two problems: unlike the emerging pattern in the Squares Problem, which is true, the emerging pattern in the Circle and Spots Problem fails for $n = 6$. Our plan was for the teacher to use the anticipated cognitive conflict that students would experience with the unexpected discovery of the counter-example in order to help them move from the naïve empirical to the crucial experiment conception.

After about 10 minutes of small group work the teacher brought the class together and asked the students whether they thought they had an answer for $n = 15$. Mac said his group thought the formula included powers of 2. The teacher asked the class to state the maximum number of non-overlapping regions they found for different numbers of spots, and drew a table on the board with the following numbers: 4, 8, and 16, for $n = 3, 4$, and 5 respectively. Then she pointed out to the class that as Mac had mentioned, the values were all powers

Place different numbers of spots around a circle and join each pair of spots by straight lines. Explore a possible relation between the number of spots and the maximum number of non-overlapping regions into which the circle can be divided.

When there are 15 spots around the circle, is there an easy way to tell for **sure** what is the maximum number of non-overlapping regions into which the circle can be divided?

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

**FIGURE 2:** The Circles and Spots Problem.

of 2; in each case, the power was one less than the number of spots: $2^2$ for $n = 3$, $2^3$ for $n = 4$, and $2^4$ for $n = 5$. The teacher asked: '*So what will it be for 15 spots, then?*'

Several students offered to answer the teacher's question. Based on what I observed during students' small group work, I presumed they would propose application of the $2^{n-1}$ formula for $n = 15$. However:

Ken:      [*said loudly*] '*Can I just say that is wrong, because on 6 [spots] there are only 30 [regions].*'

Teacher:  '*We were about to say that the answer would be 2 to the power of 14. However, you are telling me that for 6 spots it doesn't work out to be ... With this pattern for 6 six spots it would be 2 to the power of 5, that would be 32, but did anyone manage to find this number of spots?*' [*Some students said they found 31.*]

'*When we were back to the Squares Problem, we said that because the pattern worked for some of the different grids, the 5-by-5, 6-by-6 squares, and so on, we were willing to trust it. But this time we have shown that it works for 3, it works for 4, it works for 5, but actually, Ken, you are right: if we had 6 spots on a circle and we joined them all up, the number of non-overlapping regions that we get is not what we expect to get, it's not 32. It's actually 31.*'

(Stylianides, 2010a, pp. 12–13)

As she talked, the teacher used a PowerPoint slide to illustrate a case in which the maximum number of non-overlapping regions for $n = 6$ was 31, noting also that this number of regions was the maximum one could get with 6 spots.[3] She also noted that if one drew the spots in a regular hexagon, the maximum number of regions would be 30, which is again smaller than 32. Then, following the lesson plan, she asked the students to write down their thoughts on what the Circle and Spots Problem had taught them. The same three students as quoted before wrote as follows:

You can't always trust a formula until you have tested it many times over for lots of different examples [Bob].

This test has taught us that if you see a pattern [*it*] doesn't make it correct [Calvin].

The circle and spots tells us that we can't always trust a formula that works on the first few [Dan].

(Stylianides, 2010a, p. 13)

As suggested by these reflections, which were again representative of the whole class, the students had started to move away from naïve empiricism. For example, Dan started feeling uneasy trusting a pattern based on checks of the first few cases. Also, Bob's comment approximated the crucial experiment conception, since he seemed to raise a concern about the number and nature of cases one had to check before trusting a pattern.

3. The question in the Circle and Spots Problem asked whether there is an *easy way* to tell *for sure* what the maximum number of non-overlapping regions into which the circle can be divided is for $n = 15$ (see Figure 2). Although the students' inability to generate 32 regions for $n = 6$ did not guarantee that it was impossible to generate this number of regions with 6 spots, it did suggest that the emerging pattern offered an *insecure* way to find the maximum number of regions for $n = 15$. In this sense, the teacher's explanation to the class that the maximum number of regions for $n = 6$ was 31 (which is true) should be interpreted as a confirmation of students' emerging view that the pattern they identified for $n \leq 5$ offered an untrustworthy means to answer the question for $n = 15$, rather than as an authoritative act to impose conviction of a certain truth in her class. After all, the class did not have the means to prove the complex formula that describes the correct pattern.

Indeed, an important issue for many students at this stage of the lesson was how many cases would be enough for them to check before trusting a pattern. This issue had been anticipated in our planning, and we prepared a PowerPoint slide with the following fictitious student comment, that the teacher used to orchestrate discussion:

> The Circle and Spots Problem teaches me that checking 5 cases is not enough to trust a pattern in a problem. Next time I work with a pattern problem, I'll check more cases to be sure.

(Stylianides, 2010b, slide 1)

The teacher invited reactions to this comment from the class. As examples of responses, Dan suggested trying spread cases, such as for 1 spot, 75 spots and 100 spots, and Larry said: '*You should test it* [*the formula*] *as many times as you have time to do.*' The teacher asked Larry: '*So when you have tested it as many times as you have time to do, can you then trust it?*' Larry revised his response: '*No … not 100 per cent!*' Then Pauline said: '*Try it out with smaller numbers and bigger numbers.*' The teacher observed that Pauline's comment was similar to Dan's earlier one. Indeed, the two comments were similar and illustrative of the crucial experiment conception, which predominated in the class at this stage of the intervention.

### Activity 3: The Monstrous Counter-example Illustration

The teacher introduced the Monstrous Counter-example Illustration[4] (Figure 3) on a PowerPoint slide, which she presented in segments to give to the students a chance to process the information. Once the students were comfortable with the meaning of the statement, the teacher showed the counter-example. The students were amazed and experienced a new cognitive conflict, for they had not anticipated that a pattern that held for so many cases (of the order of septillions) could ultimately fail.

The teacher then directed the students' attention to their previous discussion about the fictitious student comment:

Teacher: '*We said in the Circle and Spots Problem that, okay, it's not enough to just check a few cases, you need to try different ones. Well, this expression, what does this tell us?*'

Emily: '*If you kept trying, you might have to go that high until you find one* [*a counter-example*].'

Teacher: '*But I can imagine that it took the computer quite a long time to check all of those cases. And when do you stop checking?*'

Larry: '*When you've found one!*' [*Several students laughed.*]

Teacher: '*And when do you trust a pattern then?*'

Adam: '*When you cannot find one, until you are dead!*'

(Stylianides, 2010a, p. 15)

The previous discussion illustrates that the students began to develop distrust in empirical arguments of any kind, including crucial experiment, and proposed checking the pattern indefinitely. Yet whilst students began to realise the limitations of empirical arguments, they lacked knowledge of more secure methods of validation. This caused frustration in some, as illustrated in Adam's comment to the effect that one could die whilst checking cases before being in a

--------

4.The name 'Monstrous Counter-example' was not mentioned in the class; this name is used in the article for ease of reference to the illustration.

Consider the following statement:

The expression
$$1 + 1141n^2, \text{ where } n \text{ is a natural number}$$
*never* gives a square number.

People used computers to check this expression and found out that it does *not* give a square number for any natural number from 1 to 30 693 385 322 765 657 197 397 207.

BUT

It *gives* a square number for the next natural number!!!

*Source*: Adapted from Davis, P.J. (1981). Are there coincidences in mathematics? *The American Mathematical Monthly, 88*, 311–320. doi:10.2307/2320105

**FIGURE 3:** The Monstrous Counter-example Illustration.

position to trust a pattern. Thus we may say that the students reached a point where they saw an intellectual need to learn about more secure validation methods, that is, they started to reason within the realm of the nonempirical conception, which was the intended endpoint of the intervention. It is beyond the scope of this article to give details about what happened next in the class.

### Ethical considerations

The data for this article were collected when I held an academic fellowship at the University of Oxford. The data collection and broader study were scrutinised according to the University's procedures for ethical approval, which met the British Educational Research Association and British Psychological Society standards. School, teacher, and student participation in the study was voluntary and was solicited through consent forms to the school's head teacher, the school's head of mathematics, the teachers, their students, and the students' parents or guardians. All student names mentioned in the description of the instructional intervention were pseudonyms.

### What knowledge was important for the teacher to have to successfully implement the instructional intervention in her classroom?

I return now to the question I invited the readers to consider whilst reading the description of implementation of the instructional intervention. The elements of mathematical knowledge about proof and those of knowledge about students' conceptions of proof discussed previously were necessary for successful implementation of the instructional intervention. Specifically, if the teacher had difficulty in understanding the idea that empirical arguments are not proofs (cf. mathematical knowledge about proof), she would most likely have accepted students' empirical arguments as proofs in the Squares Problem, and would have seen no purpose in proceeding with the rest of the instructional intervention. Similarly, if the teacher did not understand the taxonomy of student conceptions of proof and how its different components compare in terms of mathematical sophistication (cf. knowledge about students' conceptions of proof), she would have had difficulty evaluating her students' thinking at different stages during the intervention. For example, without this knowledge she might not have recognised the crucial experiment approach to validation that

surfaced in students' work on the Circle and Spots Problem as a significant progression in students' learning, compared to the naïve empirical approach that dominated their earlier work on the Squares Problem.

Despite the importance of the teacher's mathematical knowledge about proof and about students' conceptions of proof, as explained earlier, these two kinds of knowledge are by themselves inadequate to capture the breadth of knowledge that she used (consciously or not) during implementation of the intervention. The teacher also drew on a strong pedagogical knowledge, some elements of which were generic and characteristic of her broader teaching practice: for example, her good questioning strategies and ability to create and maintain a positive classroom environment in which the students felt comfortable contributing and challenging ideas, making mistakes, and so on. In addition to these generic elements, however, the pedagogical knowledge which she exhibited included at least two key elements that were specific to the intervention.

These two key elements, which presumably would be required by other teachers who might successfully implement the intervention in their classrooms, are discussed next.

### Element 1: Understanding how the three activities fitted together to form a coherent whole and to support students' progression along the intended learning trajectory

The three activities formed a coherent whole that supported students' progression along the intended learning trajectory. Research on prior iterations of the instructional intervention showed that possible changes in organisation of the activities in the intervention are unlikely to have such a powerful and positive effect on students' learning (see Stylianides & Stylianides, 2009, pp. 331–333). For example, this research suggested that if the teacher omitted the Circle and Spots Problem and went directly from the Squares Problem to the Monstrous Counter-example Illustration, many students who at the time held the naïve empirical conception would treat the Monstrous Counter-example in the illustration as an exception, that is, as an extreme case outside of their experience. Accordingly, the counter-example in the illustration would fail to become 'pivotal' for students (see Zazkis & Chernoff, 2008); that is, it would fail to generate a cognitive conflict for them, thereby having no substantial effect on their naïve empirical conceptions.

More generally, a teacher would need to understand the overarching idea which permeated the choice and sequencing of activities in the intervention, namely to confront students with a series of cognitive conflicts, each linked to a pivotal counter-example and intended to facilitate students' progression to the next conception in the intended learning trajectory. The counter-example in the Circle and Spots Problem corresponded to a case that was immediately outside of the convenient set of cases that (based on past experience) students with the naïve empirical conception would check in this problem ($n \leq 5$). As a result, the discovery of the counter-example (for $n = 6$) had good potential to challenge (as it did) students' naïve empirical conception, thereby facilitating their progression to the crucial experiment conception.

At that stage students considered, as we saw in the previous section, that more strategically selected checks would help discover possible counter-examples to pattern problems. In other words, they felt that crucial experiment would offer a secure method of validation. Students' crucial experiment conception was then challenged by the counter-example in the Monstrous Counter-example Illustration. This counter-example corresponded to a remarkably large case that one would not normally expect to discover even if one strategically selected which cases to check. The new counter-example was pivotal for the students and facilitated their further progression to the nonempirical conception.

To conclude, a teacher would have to understand that possible changes in the organisation of activities in the intervention would probably disrupt the intended series of cognitive conflicts, thereby having a negative effect on the potential of the intervention to support the intended learning trajectory. The choice and sequencing of activities in the intervention were purposeful and based on an empirically tested theoretical framework for how instruction can generate cognitive conflicts for students in the area of proof by means of pivotal counter-examples (see Stylianides & Stylianides, 2009, pp. 319–323).

### Element 2: Understanding the nuances of implementing the activities – The notion of 'conceptual awareness pillars'

Understanding the organisation of the activities in the sequence is important, but is nevertheless inadequate to support successful implementation of the intervention by itself. Successful implementation also requires that teachers understand the nuances of implementing the activities and, in particular, the following important premise that underpinned their design. Students are more likely to experience a cognitive conflict in the area of proof when there is a sharp contrast between their existing conceptions of proof and a situation that contradicts these conceptions: the more aware students are of their existing conceptions, the more likely it is that they will experience a cognitive conflict when they encounter a counter-example intended to challenge these conceptions (Stylianides & Stylianides, 2009).

Specifically, in order to help the students become more aware of their conceptions of proof at different stages during the intervention, the teacher used (following the lesson plans) four 'conceptual awareness pillars', that is, 'instructional activities that aim[ed] to direct students' attention to their conceptions about a particular mathematical topic' (Stylianides & Stylianides, 2009, p. 322), in this case proof. The teacher used conceptual awareness pillars both before and after the two counter-examples, which were intended to generate cognitive conflicts for students. The conceptual awareness pillars that came before the counter-examples aimed to prepare students for the upcoming cognitive conflicts, whilst those that followed them aimed to focus students' attention on the validation issues raised by the counter-examples and on thinking about possible ways to resolve the cognitive conflicts.

The first conceptual awareness pillar was at the end of the Squares Problem, when the teacher asked the students to respond to the prompt about whether and why they could be sure that applying the formula $(n - 2)^2$ for $n = 60$ would indeed give the correct answer to part three of the Squares Problem. By each student writing down their reasons for their certainty about the correctness of the formula, the students became more aware of their validation methods at the time (predominantly naïve empirical).

The second conceptual awareness pillar was again at the end of the Squares Problem, when the teacher reported back to the students the dominant (most common) validation method in the class (according to her analysis of their individual responses to the prompt in the first conceptual awareness pillar). With this report, the teacher helped the students become even more aware of the validation method that predominated in the class at the time, thereby preparing them to experience a cognitive conflict when they encountered the counter-example in the Circle and Spots Problem.

The third conceptual awareness pillar followed the counter-example in the Circle and Spots Problem, and took the form of a class discussion around the fictitious student comment on how many cases would suffice to be checked before one could trust a pattern in mathematics. By orchestrating a discussion around this fictitious comment, the teacher helped direct students' attention to the implications of the counter-example in the Circle and Spots Problem for their conceptions about validating patterns in mathematics. As shown in the earlier description of implementation of the intervention, this discussion resulted in several students resolving the prior cognitive conflict by progressing to the crucial experiment conception. By articulating their ideas of what it meant for them to validate a pattern in mathematics, the students became more aware of their new conceptions of proof, thereby reaching a state of mind amenable to experiencing (as intended) a new cognitive conflict in the context of the Monstrous Counter-example Illustration.

The fourth (last) conceptual awareness pillar followed the presentation of the Monstrous Counter-example Illustration, and took the form of a new class discussion around the same fictitious student comment as before. By orchestrating this discussion, the teacher directed her students' attention to issues of validation as opposed to other possible issues they could consider that would not be directly relevant to the goals of the lessons (e.g. the issue of using computers in mathematics). In addition, the teacher used the discussion to focus her students' thinking on the implications of the Monstrous Counter-example Illustration for the veracity of the conceptions of proof they articulated in their earlier discussion of the fictitious student comment (predominantly in the form of crucial experiment). Consideration of these implications and the ways in which they could be resolved resulted in the class progressing to the nonempirical conception of proof.

To conclude, a teacher would have to understand the rationale for including all four conceptual awareness pillars in the intervention, so that she or he could implement them in the classroom without jeopardising their effect on students' learning experience. Research on prior iterations of the instructional intervention showed that possible omission of some or all of these conceptual awareness pillars is likely to weaken the potential the two counter-examples in the intervention have to be pivotal for students and to create the intended cognitive conflicts (Stylianides & Stylianides, 2009).

## Conclusion

This article focused on the common and deeply rooted student misconception that empirical arguments are proofs, and aimed to contribute to development of a more comprehensive knowledge package for teaching proof. Specifically, the importance of expanding teachers' knowledge to include, in addition to knowledge about mathematics and students, fine-grained, domain-specific pedagogical knowledge that can allow them to help students overcome this misconception was highlighted. The implementation of a research-based instructional intervention found to be effective in helping secondary students begin to overcome this misconception was discussed, and two key elements of pedagogical knowledge for teaching proof that are important for teachers to have when they implement the intervention in their classrooms were identified.

Although the two elements of pedagogical knowledge identified are specific to the focal intervention, the notions of *pivotal counter-examples* and *conceptual awareness pillars* that underpin these elements are generic and could very well find application in other instructional interventions, not only in the area of proof but also in other areas of the school curriculum where students have persistent misconceptions. It is also important to recognise that these two elements of pedagogical knowledge are based (implicitly or explicitly) on certain premises about how students learn mathematics best, and how teaching can support that learning. For example, a premise that underpinned both elements (as well as the design of the instructional intervention as a whole) was that deeply rooted student misconceptions cannot be changed simply by 'telling' from the teacher, but rather by purposeful didactical engineering and careful design of activities that can drive students through a series of cognitive conflicts. Thus, a different perspective on teaching and learning could lead to the design of a different instructional intervention, which in turn could require different elements of pedagogical knowledge from the teacher. Yet, the limited progress that teaching practice and research have made thus far to address the pervasiveness amongst students of the misconception that empirical arguments are proofs elevates the merit of the specific instructional intervention, which has opened a window of optimism by showing how it might be possible to attack this stubborn problem in students' mathematical education.

Students tend to have several other deeply rooted misconceptions in the area of proof, and this creates many challenges for educators in designing instructional interventions to successfully address them. Indeed, it took us five research cycles of design, implementation, analysis and refinement over a four-year period (Stylianides & Stylianides,

2009) before we managed to theorise and develop the instructional intervention, a modified version of which was presented in this article. It would therefore be unrealistic to expect teachers to develop such interventions on their own, and also the knowledge about how they could successfully implement the interventions in their classrooms. How can the field of mathematics teacher education support teachers in promoting student learning of proof?

One way would be to make 'educative curriculum materials' (Davis & Krajcik, 2005) available to teachers, which can incorporate existing research knowledge about instructional interventions for promoting student learning of proof, like that discussed here. By definition, educative curriculum materials would not only present the activities in the instructional interventions, but also help teachers develop the kinds of knowledge (about mathematics, students and pedagogy) that could allow them to implement effectively these activities in the classroom. However, despite the important role that educative curriculum materials can play in promoting teacher knowledge for teaching proof, there is still a long way to go in the development of such materials. For example, research has shown that a popular reform-oriented textbook series in the United States offered limited support to teachers about how to implement the proof tasks included in the series in their classrooms (Stylianides, 2007c).[5]

Finally, another way in which the field could support teachers to acquire important knowledge for teaching proof would be to integrate into teacher preparation programmes a coherent set of learning experiences for prospective teachers, that: (1) would address all elements of a comprehensive knowledge package for teaching proof; and (2) would also consider prospective teachers' affective worlds, notably their beliefs about teaching proof. With regard to the former, we saw, for example, that various elements of the specific knowledge package discussed here are interconnected, and therefore that possible consideration of only some of these elements is likely to be insufficient to prepare teachers for successful implementation of the respective instructional intervention. With regard to the latter, it is important to help prospective teachers develop beliefs about proof that are consistent with desirable school-based instructional practices and objectives for teaching proof to students (e.g. National Council of Teachers of Mathematics, 2000). Incompatible teacher beliefs are likely to hinder development of these practices and promotion of these objectives (Thompson, 1984).

## Acknowledgements

### Competing interests

5.This textbook series has recently undergone revision, and it is possible that the issues identified by Stylianides (2007c), whose examination focused on the earlier version, have now been addressed.

# References

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

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

Ball, D.L., & Bass, H. (2003). Making mathematics reasonable in school. In J. Kilpatrick, W.G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 27–44). Reston, VA: National Council of Teachers of Mathematics.

Ball, D.L., Thames, M.H., & Phelps, G. (2008). Content knowledge for teaching. What makes it special? *Journal of Teacher Education, 59*, 389–407. doi:10.1177/0022487108324554

Chazan, D. (1993). High school geometry students' justification for their views of empirical evidence and mathematical proof. *Educational Studies in Mathematics, 24*, 359–387. doi:10.1007/BF01273371

Clements, D.H., & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning, 6*(2), 81–89. doi:10.1207/s15327833mtl0602_1

Coe, R., & Ruthven, K. (1994). Proof practices and constructs of advanced mathematics students. *British Educational Research Journal, 20,* 41–53. doi:10.1080/0141192940200105

Davis, B., & Simmt, E. (2006). Mathematics-for-teaching: An ongoing investigation of the mathematics that teachers (need to) know. *Educational Studies in Mathematics, 61,* 293–319. doi:10.1007/s10649-006-2372-4

Davis, E.A., & Krajcik, J.S. (2005). Designing educative curriculum materials to promote teacher learning. *Educational Researcher, 34*(3), 3–14. doi:10.3102/0013189X034003003

Davis, P.J. (1981). Are there coincidences in mathematics? *The American Mathematical Monthly, 88*, 311–320. doi:10.2307/2320105

Dewey, J. (1903). The psychological and the logical in teaching geometry. *Educational Review, XXV*, 387–399.

Goulding, M., & Suggate, J. (2001). Opening a can of worms: Investigating primary teachers' subject knowledge in mathematics. *Mathematics Education Review, 13,* 41–54. Retrieved from http://www.ametonline.org.uk/index.php?option=com_docman&task=doc_download&gid=18&Itemid=12

Goulding, M., Rowland, T., & Barber, P. (2002). Does it matter? Primary teacher trainees' subject knowledge in mathematics. *British Educational Research Journal, 28*(5), 689–704. doi:10.1080/0141192022000015543a

Hanna, G. (2000). Proof, explanation and exploration: An overview. *Educational Studies in Mathematics, 44*, 5–23. doi:10.1023/A:1012737223465

Harel, G. (1998). Two dual assertions: The first on learning and the second on teaching (or vice versa). *The American Mathematical Monthly, 105,* 497–507. doi:10.2307/2589401

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

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education, 31*, 396–428. doi:10.2307/749651

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

Küchemann, D., & Hoyles, C. (2001–2003). *Longitudinal proof project (technical reports for Year 8, 9, and 10 surveys).* London: Institute of Education.

Ma, L. (1999). *Knowing and teaching elementary mathematics.* Mahwah, NJ: Lawrence Erlbaum.

Martin, W.G., & Harel, G. (1989). Proof frames of preservice elementary teachers. *Journal for Research in Mathematics Education, 20*, 41–51. doi:10.2307/749097

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

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

Schoenfeld, A.H. (2006). Design experiments. In J.L. Green, G. Camilli, P.B. Elmore, A. Skukauskaite, & E. Grace (Eds.), *Handbook of complementary methods in education research* (pp. 193–205). Washington, DC: American Educational Research Association.

Senk, S.L. (1985). How well do students write geometry proofs? *The Mathematics Teacher, 78*(6), 448–456.

Shulman, L.S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4–14. doi:10.3102/0013189X015002004, doi:10.2307/1175860

Shulman, L.S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1–22.

Simon, M.A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education, 26*, 114–145. doi:10.2307/749205

Sowder, L., & Harel, G. (1998). Types of students' justifications. *The Mathematics Teacher, 91*, 670–675.

Stylianides, A.J. (2007a). Introducing young children to the role of assumptions in proving. *Mathematical Thinking and Learning, 9*, 361–385. doi:10.1080/10986060701533805

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

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

Stylianides, A.J. (2009b). Towards a more comprehensive "knowledge package" for teaching proof. In J.H. Meyer & A. van Biljon (Eds.), *Proceedings of the 15th Annual Congress of the Association for Mathematics Education of South Africa,* Vol. 1 (pp. 242–263). Bloemfontein: Association for Mathematics Education of South Africa. Available from http://www.amesa.org.za/amesa2009/volume109.pdf

Stylianides, A.J. (2010a). *Dataset produced by the ESRC-funded project (RES000222536A): Enhancing students' proof competencies in secondary mathematics classrooms (Main dataset file)*. UKDA-store. Available from http://store.data-archive.ac.uk/store/

Stylianides, A.J. (2010b). *Dataset produced by the ESRC-funded project (RES000222536A): Enhancing students' proof competencies in secondary mathematics classrooms (Appendix 1c)*. UKDA-store. Available from http://store.data-archive.ac.uk/store/

Stylianides, A.J., & Al-Murani, T. (2010). Can a proof and a counterexample coexist? Students' conceptions about the relationship between proof and refutation. *Research in Mathematics Education, 12*(1), 21–36. doi:10.1080/14794800903569774

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. doi:10.1007/s10857-008-9077-9

Stylianides, G.J. (2007c). Investigating the guidance offered to teachers in curriculum materials: The case of proof in mathematics. *International Journal of Science and Mathematics Education, 6*, 191–215. doi:10.1007/s10763-007-9074-y

Stylianides, G.J. (2008). An analytic framework of reasoning-and-proving. *For the Learning of Mathematics, 28*(1), 9–16. Available from http://www.jstor.org/stable/40248592

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. doi:10.1080/10986060701854425

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

Thompson, A. (1984). The relationship of teachers' conceptions of mathematics and mathematics teaching to instructional practices. *Educational Studies in Mathematics, 15*, 105–127. doi:10.1007/BF00305892

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27,* 458–477. doi:10.2307/749877

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.

Zazkis, R., & Chernoff, E.J. (2008). What makes a counterexample exemplary? *Educational Studies in Mathematics, 68,* 195–208. doi:10.1007/s10649-007-9110-4