This is a pre-publication version of:


Note: this was an invited plenary lecture
Towards a More Comprehensive “Knowledge Package” for Teaching Proof
Andreas J. Stylianides
University of Cambridge, U.K.

The concept of proof is central to meaningful learning of mathematics, but is hard for students to learn. What knowledge might allow teachers to effectively teach proof to their students? Existing research has constructed a significant, albeit incomplete, “knowledge package” for teaching proof. I summarize major elements of this package and contribute to its further development by focusing on elements of knowledge that have not received much attention and relate to teachers’ ability to implement successfully in their classrooms instructional interventions that address major student misconceptions. I exemplify this expanded knowledge package in the context of a research-based instructional intervention that aimed to help students begin to overcome the common misconception that empirical arguments are proofs.

Empirical Arguments vs Proofs
Consider the generalization: “The sum of any two odd numbers is an even number.” What argument would your students offer for it? Would that be a proof?

An overwhelming body of research shows that students of all levels of schooling, including high-attaining secondary students, “prove” mathematical generalizations such as the above by using empirical arguments (e.g., Coe & Ruthven, 1994; Healy & Hoyles, 2000; Senk, 1985). By empirical arguments I mean arguments that purport to show the truth of a generalization by validating the generalization in a proper subset of all possible cases (see Balacheff, 1988; Sowder & Harel, 1998; Stylianides, 2008). Empirical arguments are invalid, because they cannot exclude the possibility of the existence of a counterexample to the generalization. Here are two examples of empirical arguments for the above generalization:

Empirical argument 1: naïve empiricism
I tried many different pairs of odd numbers and their sum was always an even number: $7 + 9 = 16$, $15 + 21 = 36$, $25 + 27 = 52$, etc. So the sum of any two odd numbers is an even number.

Empirical argument 2: crucial experiment
I checked different kinds of pairs of odd numbers: some with small odd numbers (e.g., $1 + 9 = 10$), some with big odd numbers (e.g., $213 + 399 = 612$), some with the same odd numbers (e.g., $25 + 25 = 50$), and some with prime odd numbers (e.g., $17 + 31 = 48$). No pair gave me a counterexample – the sum was always an even number. So the sum of any two odd numbers is an even number.
Even though both arguments are invalid, the second argument can be considered more advanced than the first, because the search of possible counterexamples in the second involves a strategic selection of cases in contrast to the random (or convenience) sampling of cases in the first. Balacheff (1998) used the terms naïve empiricism and crucial experiment to describe the special kinds of empirical arguments represented by the first and second examples, respectively.

The fact that a generalization is true in some cases does not guarantee, and thus does not prove, that the generalization is true for all possible cases. This is the main limitation of any kind of empirical argument, but many students do not understand it. What would, then, be a proof for the generalization? Figure 1 shows three possible proofs for the generalization on the set of whole numbers.

Figure 1: Three possible proofs (on the set of whole numbers) for “odd + odd = even” (Stylianiides & Stylianides, 2008, p. 108).

Notice the correspondences among the three arguments: they are saying the “same thing” using different representations. Notice also how each argument can be used to not only help someone understand why the generalization is true, but also to convince someone that the generalization is true for all cases without requiring that person to make a leap of faith. A proof’s potential to promote understanding (explanation) and conviction (justification) accounts in part for why proof is considered to be fundamental to meaningful learning in mathematics (see, e.g., Ball & Bass, 2000, 2003; Hanna, 2000; Harel & Sowder, 2007; Mason, 1982; Stylianides & Stylianides, 2008). 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).

Unless students realize the limitations of empirical arguments as methods for validating generalizations, they are unlikely to appreciate the importance of proof in
mathematics. For this learning objective to be achieved, however, it is necessary that teachers have good knowledge in the domain of proof: the quality of learning opportunities students receive in classrooms depends on the quality of their teachers’ knowledge (e.g., Ball, Thames, & Phelps, 2008; Ma, 1999; Shulman, 1986). What knowledge might allow teachers to effectively help their students begin to overcome the misconception that empirical arguments count as proofs?

**KNOWLEDGE FOR TEACHING PROOF**

Effective teaching of proof requires, at a minimum, three broad kinds of teacher knowledge:

- *Mathematical knowledge about proof,* i.e., a solid understanding of critical mathematical aspects of proof that are essential for teaching proof to students;
- *Knowledge about students’ conceptions of proof,* i.e., a solid understanding of common ways in which students think about proof (including misconceptions); and
- *Pedagogical knowledge about proof,* i.e., a good command of effective pedagogical practices for helping students develop conceptions of proof that better approximate conventional understandings.

Next I summarize what the field of mathematics education currently considers important for teachers to know in relation to these three broad kinds of knowledge, paying particular attention to the distinction between empirical arguments and proofs. Then I use this summary to discuss important aspects of research knowledge that is currently missing for the development of a more comprehensive “knowledge package” for teaching proof. My use of the term “knowledge package” differs from that of Ma (1999). Ma used this term to describe complex networks of relationships among 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 have to (or can) be learned simultaneously, etc. Thus, Ma used knowledge packages to describe concept maps or organizing structures within teachers’ mathematical knowledge for teaching. My use of the term is broader than that. I use the term knowledge package to describe a cluster of related kinds of knowledge (notably, knowledge about mathematics, students, and pedagogy) that a teacher must have in order to teach effectively a particular concept. Accordingly, the knowledge package for teaching proof can include, but is not limited to, a concept map that shows the relationships between proof and other closely connected concepts such as pattern, conjecture, and argument.

**Mathematical Knowledge about Proof**

In Stylianides and Ball (2008), we conducted a comprehensive review of the literature on teachers’ mathematical knowledge about proof for teaching and we
contributed to the further development of this literature. An important element of mathematical knowledge about proof that we identified in that article, which is of interest to this article, is that teachers need to understand the distinction between empirical arguments and proofs (as explained in first section of the article). Unless teachers of all levels of education develop good understanding of the distinction between empirical arguments and proofs, it is unlikely that large numbers of students will overcome the misconception that “empirical evidence = proof.” For example, Martin and Harel (1989) noted about elementary teachers: “If [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” (pp. 41-42). Furthermore, an elementary teaching practice that promotes or tolerates a conception of proof as an empirical argument instills inaccurate mental habits in students. Dewey (1903) cautioned educators against such practices: he said that, whatever the preliminary approach to learning is, 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” (p. 217).

Knowledge about Students’ Conceptions of Proof

A significant body of research investigated students’ conceptions of proof and has developed various taxonomies of students’ conceptions of proof (some of which have developmental nature) that inform the field’s understanding of what is important for teachers to know in this area (see, e.g., Balacheff, 1988; Sowder & Harel, 1998). Good knowledge about students’ conceptions of proof can help teachers to identify/describe their own students’ understandings about proof. Also, this knowledge can inform the design of instruction that will aim to help students develop conceptions of proof that better approximate conventional understandings.

Below is a taxonomy of students’ conceptions of proof presented in increasing level of mathematical sophistication (see Stylianides and Stylianides [2009] for elaboration on this taxonomy).

- **Conception 1**: validating mathematical generalizations (e.g., patterns, conjectures) using naïve empiricism;
- **Conception 2**: validating mathematical generalizations using crucial experiment;
- **Conception 3**: recognizing empirical arguments as insecure methods for validating mathematical generalizations (i.e., seeing a need to learn about more secure methods of validation); and
- **Conception 4**: recognizing and using proofs as secure methods for validating mathematical generalizations.
As I noted earlier, naïve empiricism and crucial experiment are two kinds of empirical arguments (Balacheff, 1988), with the former kind being a less advanced version of the latter. So, although conceptions 1 and 2 are both representative of an empirical approach to validating mathematical generalizations, conception 1 is less advanced than conception 2. These two conceptions are dominant among students of all levels of schooling, including high-attaining secondary students (e.g., Coe & Ruthven, 1994; Healy & Hoyles, 2000; Senk, 1985). Conception 3 is a transitory conception between the empirical approach to validating mathematical generalizations (conceptions 1 and 2) and the conventional approach (conception 4).

**Pedagogical Knowledge about Proof**

Compared to the body of research that informed the previous two kinds of teacher knowledge, significantly less research attention has been paid to the identification of pedagogical practices that teachers need to know for effective teaching of proof. Existing research in this area has developed a useful research base about general pedagogical practices for engaging students in mathematical reasoning, argumentation, and proof (see, e.g., Ball & Bass, 2000, 2003; Stylianides, 2007a, b; Stylianides & Ball, 2008; Yackel & Cobb, 1996; Zack, 1997). Yet, this research base needs to be developed further before it can inform the design of effective instructional interventions that will help students develop more accurate conceptions of proof; whereby “instructional intervention” refers to a purposeful and cohesive collection of activities, and respective implementation strategies of these activities, for achieving particular learning outcomes.

For example, in Stylianides and Ball (2008) we elaborated the importance of teachers using a variety of proving tasks that can offer students learning opportunities to develop understanding of different proving strategies and reasoning skills. Although the classroom implementation of particular proving tasks can provoke construction of different kinds of arguments, thereby offering to teachers and students the opportunity to discuss and reflect on the differences between empirical arguments and proofs, it is unclear how this discussion/reflection can be organized in a way that will help students overcome their deeply rooted misconception that “empirical evidence = proof.” Indeed, research and practice showed that addressing this misconception is a stubborn problem in mathematics education (see, e.g., Goulding & Suggate, 2001), and so this problem cannot be resolved without a carefully designed instructional intervention. Ability to implement successfully such an instructional intervention would be an essential element of teachers’ pedagogical knowledge about proof and, by implication, an important complement to the developing knowledge package for teaching proof.

What might such an instructional intervention look like and what demands does the successful implementation of the intervention place on teachers’ knowledge? Next I will discuss a research-based instructional intervention that was shown to be effective in helping students begin to overcome the misconception that “empirical evidence =
My discussion will also exemplify the point that successful implementation of the instructional intervention presupposes that the teacher who implements the intervention has, in addition to good understanding of the different aspects of the intervention itself, solid knowledge of the mathematical ideas about proof and students’ conceptions of proof that I discussed earlier. The exemplification of this point will emphasize the inextricable relationships among the different kinds of knowledge that comprise the knowledge package for teaching proof.

AN INSTRUCTIONAL INTERVENTION FOR HELPING STUDENTS BEGIN TO OVERCOME THE MISCONCEPTION THAT “EMPIRICAL EVIDENCE = PROOF”

Background

In a four-year design experiment in an undergraduate mathematics course in the United States (see Schoenfeld [2006] for discussion on design experiment methodology), Gabriel Stylianides and I developed an instructional intervention that we showed to be effective in helping undergraduate students begin to understand the limitations of empirical arguments and see an “intellectual need” (Harel, 1998) for learning about more secure methods of validation (i.e., proofs) (Stylianides & Stylianides, 2009).

Next I will present how a secondary mathematics teacher implemented in her class a modified version of the original instructional intervention; the modification was done by me and was discussed with the teacher to ensure (1) the appropriateness of the activities for her class and (2) that she understood the rationale of the different activities that comprised the intervention and was well prepared to implement the intervention in her class. The implementation of the intervention was part of a school-based design experiment that I conducted last year in two high-attaining Year 10 classes in a state school in England (student ages: 14-15 years old). Partially motivated by research findings that showed that even high-attaining secondary students in England possessed limited understanding of proof (Coe & Ruthven, 1994; Healy & Hoyles, 2000; Küchemann & Hoyles, 2001-03), the design experiment aimed to generate theoretical and practical knowledge about possible instructional interventions for helping Year 10 students develop their understanding of proof. It is envisioned that the knowledge to be generated by the design experiment will support future studies that will aim to promote similar research goals in more challenging settings (notably, in classes with less advanced students).

The focal instructional intervention at the school level was underpinned by the theoretical framework that we developed in our undergraduate design experiment (Stylianides & Stylianides, 2009). The student learning outcomes from the implementation of the intervention at the undergraduate and school levels were strikingly similar. This observation, together with the fact that the misconception “empirical evidence = proof” was found to be dominant among students of all levels
of education and in different countries, offer a reasonable basis for one to hypothesize that appropriately modified versions of the instructional intervention can yield similar results when implemented in other levels of education (notably, in the elementary and middle schools) and in other cultural contexts.

The implementation of the focal instructional intervention at the school level lasted approximately 60 minutes and extended over two consecutive 45-minute lesson periods. The teacher implemented the instructional intervention following a detailed lesson plan that I prepared and discussed thoroughly with her prior to the lessons. The description of the implementation in the following sections of the article is based on video and audio records of the lessons and field notes that I took on the work of a group of six students during the lessons (there was no special reason for the selection of this small group; all small groups were considered to be of mixed ability).

As you read the description of the implementation of the instructional intervention, I invite you consider what knowledge the teacher needed to have for successful implementation of the intervention (I will address this issue after I describe the implementation of the intervention). Also, I invite you to pay attention to how the teacher used each of the activities in the intervention (Squares Problem, Circle and Spots Problem, and “Monstrous Counterexample” Illustration) to facilitate students’ progression along the “learning path” that is summarized in Figure 2: from (1) using naïve empiricism as a method for validating patterns, to (2) using crucial experiment, to (3) seeing a need to learn about more secure methods for validating patterns (i.e., to learn about proofs). Notice that the three stages in this learning path correspond to conceptions 1 through 3 that I discussed earlier under teachers’ knowledge about students’ conceptions of proof for teaching.1

![Figure 2: The three activities in the instructional intervention and the corresponding “learning path.”](image-url)
Activity 1: The Squares Problem

Kathy, the teacher, introduced the Squares Problem (Figure 3). The hardest part of the problem was the third, because it asked students to find the number of different 3-by-3 squares in a case that was difficult for them to check practically and also to explain whether and why they were sure their answer was correct.

Kathy made sure that the students understood what the problem was saying and then she asked them to work on the problem in their small groups. The small group where I was sitting during the lesson had the following six students: Bob, Calvin, Dan, Lazarus, Robert, and Sharon. These students counted squares to answer parts 1 and 2 of the problem, and then Bob asked his peers: “Have you actually got a formula?” Dan responded: “It’s the number of … it’s n minus 2, and then squared.” Sharon showed excitement and confirmed with Dan that the answer for part 1 would be 4. Robert asked how many 3-by-3 squares there were in a 60-by-60 square (part 3) and Dan used his calculator and the formula he had described earlier to find the answer: 

\[(60 - 2)^2 = 3364.\]

At some point Kathy visited the small group and the students explained their work. Kathy then asked the students whether they were sure their answer was correct. Lazarus said with confidence, “yes,” and Kathy posed a new question: “And have you thought about why you are sure?” There was no response from the students. Kathy asked the students to think about this and to write their ideas on paper.

![The “Squares” problem](image)

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?

Figure 3: The Squares Problem
(adapted from Zack, 1997).
Dan drew figures for the 4-by-4 and 5-by-5 squares showing the 3-by-3 squares in each of them. He wrote down $58^2 = 3364$ as the answer to part 3 and also the formula $(n – 2)^2$. He concluded: “We realized that if you took 2 away from the number of cubes along the top and then square the answer you will get the number of 3x3 boxes in the grid.” The other students in the small group wrote similar conclusions.

So, what has happened thus far in the small group? 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, based on this evidence, they concluded that the pattern would hold true for all values of $n$ including $n=60$. Thus the students validated the pattern on the basis of naïve empiricism (cf. Figure 2).

The whole group discussion that followed illustrated further the use of naïve empiricism in the class, as all groups answered the three parts of the problem using the formula $(n – 2)^2$. After some discussion on the meaning of the formula, Kathy asked the class whether and why they could be sure that applying this formula would give the correct answer. Emily said: “We tried it [the formula] for a 6-by-6 square and it worked for that too.” Kathy invited further comments but the students did not have anything to add to what Emily had said.

Kathy then asked the students to write down individually their thoughts: “I want to know what your feelings are about whether this [the answer to part 3] is correct or not. You may think it is correct, you may not. If you are sure, I want to learn why you are sure.” Someone asked “what if you’re not sure?” and Kathy responded: “Then put not sure, but say why you are not sure – what makes you doubt it?”

In the focal small group the students wrote:

- Bob: “Because we have found a formula and tried it against smaller squares so we can make sure that the formula is right.”
- Calvin: “I am sure that this solution works because it worked for every one we did.”
- Dan: “I am sure that the answer is correct because it has been proved for a number of smaller grids.”
- Lazarus: “I am sure that the answer is correct because it has been tested and proved correct. The pattern will continue to 60x60.”
- Robert: “I am sure it’s correct because we did a test on the 6x6 grid and it worked.”
- Sharon: “We are sure that it is right because we have tried it for a 6x6 square as well. So we assume that it would work.”

Notice that the six 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, according to the students, it would work also for $n=60$. 
Following the students’ individual reflections, Kathy proceeded with the next item in the lesson plan, which was to summarize the students’ written responses. Kathy’s summary below was based on a quick inspection of students’ written responses as she was circulating around during students’ individual reflection time. The accuracy of Kathy’s summary was confirmed by a more careful analysis of students’ responses 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 $58^2$.

The way the students responded to Kathy’s prompt was anticipated in the lesson plan: the state of student conceptions of proof in Kathy’s class as described in the summary above (predominance of naïve empiricism) was consistent with what was reported in the research literature about students’ conceptions of proof.

After Kathy’s summary, Bob asked Kathy whether the pattern was correct and Kathy said that the class would come back to this issue later, but first they would work on a couple of other activities. Indeed, according to our lesson plan the issue about the correctness of the pattern in the Squares Problem would remain tentatively unresolved. The class would revisit and resolve the issue after the students had been assisted to realize the limitations of empirical arguments (both naïve empiricism and crucial experiment). Had the issue been resolved at this point of the lesson, this would probably require a lot of “telling” by the teacher, which would be inconsistent with the teachers’ philosophy of how students learn and the theoretical framework that underpinned the design of the instructional intervention. The intention was for the students to realize the limitations of empirical arguments on their own, by experiencing and reflecting on situations where the empirical validation method was inadequate. For the readers’ information, I note that the $(n – 2)^2$ pattern was actually correct.

**Activity 2: The Circle and Spots Problem**

Kathy introduced the Circle and Spots Problem (Figure 4) and helped the students understand what the problem was saying. Specifically, she discussed with them the meaning of the terms “maximum” and “non-overlapping regions.” Also, she clarified that the phrase “around the circle” referred to the circle’s circumference and that the spots on the circumference did not have to be equidistant. Then Kathy asked the students to work on the problem in their small groups.

Notice that, similarly to part 3 of the Squares Problem, the question in the Circle and Spots Problem (yellow box in Figure 4) was asking the students to make a statement about a case that was difficult for them to check practically. In our planning, we had anticipated that the students, like they did in the Squares Problem, 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 spots). The main difference between the two problems is that the emerging pattern in the Circle and Spots Problem fails for n=6. Our plan was for Kathy to use the anticipated surprise that the students would experience with the failing pattern to help them move from naïve empiricism towards crucial experiment (cf. Figure 2).

After about 10 minutes of small group work, Kathy brought the whole class together and said: “Circulating around I think there are some people who think they know what the answer will be for 15 [spots]. Is there anyone who is willing to tell us their number of regions, what it will be for 15 spots?”

Mac said that his group thought the formula for the problem was \((n-1)^2\) but soon after he corrected himself and said that the formula included powers of 2. Kathy asked the class to say the maximum number of non-overlapping regions they found for different spots, and she constructed a table on the board with the following numbers: 4, 8, and 16, which corresponded to \(n = 3, 4, \text{ and } 5\), respectively. Then she pointed out that, as Mac had mentioned earlier, the values were all powers of 2 and that, 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). Kathy asked: “So what will it be for 15 spots, then?”

Several students offered to answer Kathy’s question. Based on what I had observed during these students’ earlier work in their small groups, I presume they would propose the application of the \(2^{n-1}\) formula for \(n=15\). However, Ken, a student sitting in a different group, said loudly: “Can I just say that is wrong, because on 6 [spots] there are only 30 [regions].” Kathy said: “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.

Kathy continued:

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.

As she talked, Kathy used a PowerPoint slide to illustrate a case in which the number of non-overlapping regions for n=6 was 31, explaining that this number of regions was the maximum number of regions one could get with 6 spots. She noted also 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, Kathy asked the students to write down their thoughts about what the Circle and Spots problem had taught them.

The students in the focal small group wrote:

• Bob: “You can’t always trust a formula until you have tested it many times over for lots of different examples.”
• Calvin: “This test has taught us that if you see a pattern doesn’t make it correct.”
• Dan: “The circle and spots tells us that we can’t always trust a formula that works on the first few.”
• Lazarus: “This teaches us that just because something works for one thing, that doesn’t mean it will work for everything.”
• Robert: “You can’t always trust a formula until you have tested many times over for lots of different no’s of spots.”
• Sharon: “You can’t always trust a formula. You shouldn’t presume it is correct because it worked for the first few.”

Notice that the students started to move away from naïve empiricism. For example, Dan, Lazarus, and Sharon started feeling uneasy to trust a pattern based on checks of the first few cases. Also, Bob and Robert’s comments approximated the crucial experiment method of validation, as they appeared to raise a concern about the number (“many”) and quality (“different”) of cases that had to be checked before a pattern could be trusted.

Thus 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. We had anticipated this issue in our planning and we prepared a PowerPoint slide with a fictional student comment on it that Kathy used in the lesson to orchestrate a discussion around the issue. The fictional student comment said:
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.

Kathy invited reactions from her students on this comment. Dan suggested trying spread cases such as for \( n = 1, 75, \) and 100. Robert observed that “you can’t always trust the formula, you have to test it.” Kathy asked Robert how many times one had to test a formula and Robert said “more than like 5 times.” Kathy invited more comments and Larry said: “you should test it as many times as you have time to do.” Kathy asked Larry: “So when you have tested it as many times as you have time to do, can you then trust it?” Larry revised: “No … not a 100%!” Then Pauline said: “try it out with smaller numbers and bigger numbers.” Kathy observed that Pauline’s comment was similar to Dan’s earlier comment.

Indeed, the two comments were similar to one another and illustrative of the crucial experiment method for validating patterns (cf. Figure 2). As I noted earlier, crucial experiment can be considered to be a more advanced method than naïve empiricism, but is still an invalid method because it does not exclude the possibility of a counterexample in a case that was not checked. Some students in the class were thinking along similar lines, as illustrated further by their responses to the following question by Kathy: “And then, do we trust it if it worked for all of those [cases, big and small ones]? S” Ilvia said in a low voice: “No, because you might have missed one.” Another student was heard to say: “You could spend your whole life and still miss one!” These students’ fear that a pattern can fail in a case that was not checked was manifested in the next activity we planned for the students.

Activity 3: The “Monstrous Counterexample” Illustration

Kathy introduced the PowerPoint slide in Figure 5 that shows what I call the “Monstrous Counterexample” Illustration. Kathy did not use this name during the lesson. The slide was presented in segments to give to the students a chance to process the information in it. For example, there was a discussion about how one would check whether a given number was a square number using a calculator. Also, the students confirmed the statement for particular values of \( n \) using their calculators.

Once the students checked many different cases and were comfortable with the meaning of the statement, Kathy presented the counterexample. The students were amazed: they had not anticipated that a pattern that held for so many cases (of the order of septillions) could ultimately fail.

Kathy then directed the students’ attention to their previous discussion about the fictional student comment on what the Circle and Spots Problem had taught them: “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 said: “If you kept trying, you might have to go that high until you find one [a counterexample].” Kathy said: “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 said: “When you’ve found one!” Several students laughed with what Larry had said. Kathy continued: “And when do you trust a pattern then?” Adam said: “When you cannot find one, until you are dead!”

**Consider the following statement:**

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

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

**BUT**

It \textit{gives} a square number for the next natural number!!!

Figure 5: The “Monstrous Counterexample” Illustration (adapted from Davis, 1981).

Notice that the students began to develop distrust in empirical arguments of any kind, including crucial experiment, because they were proposing checking the pattern indefinitely. Yet, although the students began to realize the limitations of empirical arguments, they lacked knowledge of more secure methods for validating patterns. This caused a feeling of frustration among some of them as illustrated in Adam’s comment: one would die checking cases before being in a position to trust a pattern. Thus we may say that the students reached the point when they felt a need to learn about more secure validation methods (cf. Figure 2).

Following the lesson plan, Kathy responded to this emerging student need, which was an outcome of the instructional intervention, by introducing her students to the notion of proof in mathematics and by engaging them in a discussion of possible criteria that an argument needed to fulfill in order to count as a proof in their class. Also, Kathy took the class back to the Squares Problem and she helped the students to develop a proof for the pattern they had identified earlier. More discussion of what happened next in Kathy’s class is beyond the scope of this article.

**What Knowledge did the Teacher Need to Have to Successfully Implement the Instructional Intervention?**

The elements of mathematical knowledge about proof and knowledge about students’ conceptions of proof that I discussed in previous sections were necessary for successful implementation of the instructional intervention. Specifically, if the
teacher did not understand that empirical arguments do not qualify as proofs (cf. mathematical knowledge about proof), she would most likely have accepted students’ empirical arguments as proofs in the Squares Problem and she would have seen neither a need nor a purpose to proceed with the rest of the instructional intervention (which aimed to challenge this empirical conception of proof).

Similarly, if the teacher did not know the taxonomy of student conceptions of proof in Figure 2 and the hierarchical structure of these conceptions as stages in a learning path (cf. knowledge about students’ conceptions of proof), she would most likely have difficulty making sense of her students’ conceptions about proof in the different stages of the intervention and how these conceptions compared to one another (in terms of their level of mathematical sophistication). For example, the teacher might have not recognized the crucial experiment approach to validation that surfaced in students’ work on the Circle and Spots Problem as a progression in their learning when compared to the naïve empirical approach that dominated students’ earlier work on the Squares Problem. Accordingly, the teacher would have been limited in her potential to understand the function of the different activities in the instructional intervention, which aimed to help students to develop conceptions of proof that increasingly approximated conventional understandings.

Despite the importance of the teacher’s mathematical knowledge about proof and her knowledge about students’ conceptions of proof as explained above, these two kinds of knowledge are by themselves inadequate to capture the knowledge that was used by the teacher during the implementation of the instructional intervention. Specifically, the teacher was also drawing on a strong pedagogical knowledge about proof that allowed her to understand the rationale that underpinned the design of the intervention and to implement it successfully in her class, thereby managing to help many students in her class progress along the intended learning path. The description of the implementation of the instructional intervention exemplified several elements of the teacher’s pedagogical knowledge about proof. Two important elements of this knowledge were:

- **Understanding the general structure of the intervention and how the activities that comprised the intervention could be used to support students’ progression along the main stages of the intended learning path.** The three activities in the instructional intervention constituted a coherent sequence of activities that could be used to help students to progress along the intended learning path; the value of the sequence was more than just the sum of the values of the individual activities that comprised it. For example, our prior experience (Stylianides & Stylianides, 2009) suggests that, if the teacher asked her students to prove the pattern in the Squares Problem before she introduced to them the Circles and Spots Problem, this would result in a completely different learning experience for the students: they would not see any reason to abandon their naïve empirical arguments. The beginning work on the Squares Problem was designed to provoke students’ use of naïve empiricism (as documented in
the literature) and to create a basis on which the subsequent activities in the intervention built to help the students develop increasingly accurate conceptions of proof: the Circle and Spots Problem challenged students’ reliance on naïve empiricism, while the Monstrous Counterexample Illustration challenged their reliance on crucial experiment. Once the class realized the limitations of empirical arguments, the class revisited the Squares Problem and produced a proof for the pattern. Thus, the Squares Problem did not constitute an isolated activity to be completed independently, but rather an integral part of the sequence of activities that comprised the intervention.

• Understanding the nuances of particular implementation strategies in the instructional intervention. It would be difficult for students to abandon, or actively engage in a process of refining, their current conceptions about proof, unless they became more aware of their current conceptions and recognized the limitations of these conceptions. By asking the students to do the individual reflection at the end of the Squares Problem, the teacher helped them to become more aware of the method of validation they used in that problem (naïve empiricism), thereby making it more likely that they experienced a “cognitive conflict”3 in subsequent activities where the use of this validation method proved to be problematic. Similarly, by organizing the discussions of the fictional student comment after the students encountered the counterexamples in the Circle and Spots Problem and the Monstrous Counterexample Illustration, the teacher created a purposeful learning context in which the students reflected on, and began to recognize, the limitations of naïve empiricism and crucial experiment, respectively.4

It is important to recognize that these elements of pedagogical knowledge about proof are based (implicitly or explicitly) on certain premises about how students learn mathematics and how teaching can support that learning. For example, a premise that underpins both elements, and the design of the instructional intervention as a whole, is that deeply rooted student misconceptions cannot be changed simply by “telling” from the teacher, but rather by careful design of rich and purposeful learning environments for students. Thus, a different philosophy of teaching/learning or theoretical perspective on the instructional design would probably support a different instructional intervention, which would in turn require a different pedagogical knowledge about proof from the teacher. Yet, the fact that existing research and practice have not identified thus far other promising instructional interventions for helping students to overcome the deeply rooted misconception that “empirical evidence = proof,” highlights the importance of the instructional intervention presented herein and supports the philosophical and theoretical perspectives that underpinned its design as elaborated in Stylianides and Stylianides (2009).
CONCLUSION

In this article, I explained that existing research on teachers’ knowledge for teaching proof has constructed a significant, albeit incomplete, knowledge package for teaching proof. Specifically, this body of research has identified critical mathematical aspects of proof that are essential for teaching proof to students (mathematical knowledge about proof) and has constructed a detailed map of common ways in which students think about proof (knowledge about students’ conceptions of proof). Less emphasis has been placed on teachers’ knowledge of effective pedagogical practices for helping students develop more accurate conceptions of proof (pedagogical knowledge about proof).

In an effort to contribute to the development of a more comprehensive knowledge package for teaching proof, I elaborated on the importance of expanding teachers’ pedagogical knowledge about proof to include ability to implement successfully in their classrooms instructional interventions that address major student misconceptions. I exemplified my proposal in the context of a research-based instructional intervention that aimed to help students overcome the deeply rooted misconception that “empirical evidence = proof.” My discussion of the implementation of the instructional intervention by a secondary mathematics teacher illustrated also the inextricable relationships among the different kinds of knowledge of the knowledge package for teaching proof that I discussed: if the teacher lacked solid knowledge of any of these kinds, she would not have been able to implement the intervention the way she did (i.e., faithfully to the plan and with good results). Furthermore, given the limited progress that practice and research have made thus far to address the pervasiveness of the misconception that “empirical evidence = proof” among students, my discussion of the implementation of the instructional intervention by the secondary mathematics teacher sends the optimistic message that, with the necessary support, typical teachers can make significant progress in addressing this stubborn problem in their classrooms.

The development of a more comprehensive knowledge package for teaching proof, coupled with a support system for helping teachers (both prospective and practicing) to incorporate the elements of this package into their own knowledge, promise major advancements in the teaching of proof and, by implication, in students’ opportunities to develop competency in proof. Yet, a lot of challenges emerge for research and practice. The deeply rooted misconceptions that many students tend to have in the domain of proof make extremely difficult the design and implementation of instructional interventions that can successfully address these misconceptions. Indeed, it took us five research cycles of implementation, analysis, and refinement over a four-year period (Stylianides & Stylianides, 2009) in order to develop the instructional intervention whose modification was presented in this article. Thus, it would be unrealistic to expect from teachers to develop such knowledge for teaching proof on their own; rather, teachers need to be offered systematic support to develop this kind of knowledge.
One way in which the field of mathematics education can support teachers to develop their knowledge for teaching proof is by making available to them “educative curriculum materials” (Davis & Krajcik, 2005), which will incorporate, for example, existing research knowledge about instructional interventions for promoting student learning of proof. Specifically, such educative curriculum materials will be concerned not only with a presentation of the activities that comprise the instructional interventions, but also with helping teachers to develop the kinds of knowledge (about mathematics, students, and pedagogy) that will allow them to effectively implement these interventions in their classrooms. Despite the important role that educative curriculum materials can play in promoting teacher knowledge for teaching proof, it is a long way to the development of such materials. For example, research showed that a popular, reform-oriented textbook series in the United States offered limited support to teachers about how to implement in their classrooms proof tasks that were included in the series (Stylianides, 2007c).

Finally, another way in which the field can support teachers to develop their knowledge for teaching proof is by promoting this knowledge in teacher preparation and professional development programs. This goal will have to be integrated into a coherent set of learning experiences for teachers that: (1) will address all components of the knowledge package for teaching proof that I discussed in this article; and (2) will aim, additionally, to develop in teachers beliefs about proof that will help them appreciate the importance of proof in their students’ mathematical education. In regard to the latter, it is important that teacher preparation and professional development programs offer to teachers opportunities to experience themselves as learners instructional interventions that aim to help them develop their own mathematical knowledge about proof (see Stylianides & Stylianides, 2009). Specifically, research on the role of teacher preparation programs in the development of prospective teachers’ beliefs that are consistent with desirable instructional practices and objectives has showed that teacher preparation courses need to engage prospective teachers in activities that reflect the demands of the mathematics classrooms in which they will someday teach (Kagan, 1992).

ACKNOWLEDGMENTS

The work reported herein received funding support from the UK’s Economic and Social Research Council (Grant Number: RES-000-22-2536) and the Spencer Foundation (Grant Numbers: 200700100 and 200800104). The opinions expressed in the article are those of the author and do not necessarily reflect the position, policies, or endorsement of either organization. Parts of this article (notably those that describe the implementation of the instructional intervention) appeared recently in a practitioners’ journal (Stylianides, 2009), but their use herein is different and serves a new purpose (to advance an argument about teachers’ knowledge for teaching proof).
NOTES

1. The teacher and student names are pseudonyms.

2. The question in the Circle and Spots Problem was asking whether there is an easy way to tell for sure what is the maximum number of non-overlapping regions in which the circle can be divided for n=15 (see Figure 4). Although the students’ inability to generate 32 regions for n=6 does not guarantee that it is impossible to generate this number of regions with 6 spots, it suggests that the emerging pattern offers an insecure way to find the maximum number of regions for n=15. In this sense, Kathy’s explanation to the class that the maximum number of regions for n=6 was less than 32 should be interpreted as a confirmation of the students’ emerging view that the pattern they identified for n≤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.

3. See Stylianides and Stylianides (2009, pp. 319-323) for a discussion of how the similar version of the instructional intervention that was implemented at the undergraduate level used “cognitive conflict” as a mechanism to support developmental progressions in students’ knowledge and how it addressed the common challenge faced by the cognitive conflict approach to mathematics teaching of students treating emerging contradictions in their knowledge as exceptions.

4. I helped the teacher to develop these two elements of knowledge in our discussions prior to the implementation of the instructional intervention in her class.

REFERENCES


