ICME-13 Monographs

Andreas J. Stylianides
Guershon Harel *Editors*

Advances in Mathematics Education Research on Proof and Proving

An International Perspective

ICME 13
Hamburg 2016

Springer

ICME-13 Monographs

Each volume in the series presents state-of-the art research on a particular topic in mathematics education and reflects the international debate as broadly as possible, while also incorporating insights into lesser-known areas of the discussion. Each volume is based on the discussions and presentations during the ICME-13 Congress and includes the best papers from one of the ICME-13 Topical Study Groups or Discussion Groups.

More information about this series at http://www.springer.com/series/15585

Andreas J. Stylianides · Guershon Harel
Editors

Advances in Mathematics Education Research on Proof and Proving

An International Perspective

Springer

Editors
Andreas J. Stylianides
Faculty of Education
University of Cambridge
Cambridge
UK

Guershon Harel
Department of Mathematics
University of California San Diego
La Jolla, CA
USA

ISSN 2520-8322 ISSN 2520-8330 (electronic)
ICME-13 Monographs
ISBN 978-3-319-70995-6 ISBN 978-3-319-70996-3 (eBook)
https://doi.org/10.1007/978-3-319-70996-3

Library of Congress Control Number: 2017958582

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

There is an international recognition of the importance of proof and proving in students' learning of mathematics at all levels of education, and of the difficulties faced by students and teachers in this area. Also, existing curriculum materials tend to offer inadequate support for classroom work in this area. All of these paint a picture of proof and proving as important but difficult to teach and hard to learn. A rapidly expanding body of research has offered important insights into this area, leading to an upsurge of publications on various aspects of proof and proving (mathematical, cognitive, social, pedagogical, philosophical, etc.) in all mathematics education research journals and in books or specialized volumes (e.g., Hanna and de Villiers 2012; Hanna et al. 2010; Reid and Knipping 2010; Stylianides 2014, 2016; Stylianides and Stylianides 2017; Stylianou et al. 2010). The state of the art in this research area has been summarized and discussed in several literature reviews (Harel and Sowder 2007; Mariotti 2006; Stylianides et al. 2016, 2017), which have shown not only the progress we have made as a field over the past few decades in addressing key questions related to proof and proving but also that there are still many open questions for which research-based responses are sorely needed.

This book explores new trends and developments in mathematics education research related to proof and proving, the implications of these trends and developments for theory and practice, and directions for future research. With contributions from researchers working in 12 different countries (Canada, Chile, France, Germany, Hong Kong, Israel, Italy, Japan, Norway, Peru, the UK, and the USA), the book brings also an international perspective to the discussion and debate of the state of the art in this important area.

The book is organized around the following four parts, which reflect the breadth of issues addressed in the book. Under each part (essentially a theme), there are four main chapters and a concluding chapter offering a commentary on the theme overall. Although several chapters addressed issues that spanned several themes, practical considerations related to the organization of the book necessitated a best-fit approach.

- Part I: Epistemological Issues Related to Proof and Proving (Chaps. 1–5);
- Part II: Classroom-based Issues Related to Proof and Proving (Chaps. 6–10);
- Part III: Cognitive and Curricular Issues Related to Proof and Proving (Chaps. 11–15);
- Part IV: Issues Related to the Use of Examples in Proof and Proving (Chaps. 16–20).

The book's main chapters (i.e., all but the four commentary chapters) are extended and revised versions of papers that were presented at Topic Study Group 18, titled "Reasoning and Proof in Mathematics Education", of the 13th International Congress on Mathematical Education (July 2016, Hamburg, Germany). Associated with this Topic Study Group, which we co-chaired, there were 68 contributions in three categories: 21 8-page papers (regular presentations), 35 4-page papers (oral communications), and 12 posters; the book's main chapters derived from the first category. These contributions passed through a rigorous, stepwise review process that included several cycles of feedback and revision. The first key step in the process was the review and subsequent extension/revision of the contributions prior to their acceptance as regular presentations at the Congress; each contribution was reviewed by at least two members of the Topic Study Group organizing team, which comprised Paolo Boero, Mikio Miyazaki, David Reid, and the two of us as co-chairs. The second key step in the process was the extension/revision of the contributions based on the feedback received during the Congress and the overall discussions that happened during the work of the Topic Study Group. The third key step in the process was the review of the post-Congress revisions by two other contributors of the book and one of us as the Handling Editor; this review led to another substantial round of extension/revision. Most chapters were accepted for inclusion in the book at the end of the third step; a few others underwent a further round of revision that was overseen by the Handling Editor. The four commentary chapters were written by invited contributors—Keith Weber and Paul Dawkins (Chap. 5), Ruhama Even (Chap. 10), Lianghuo Fan and Keith Jones (Chap. 15), and Orit Zaslavsky (Chap. 20)—and were reviewed only by us. Finally, Gabriele Kaiser, the ICME-13 Monograph Series Editor, reviewed the whole book before it was sent to production.

We wish to thank the participating authors for their dedication and cooperation; the reviewers and commentators for their diligent work; the presenters and audience of our Topic Study Group for their feedback and insight; the members of the study group organizing team, Paolo Boero, Mikio Miyazaki, and David Reid, for their valuable inputs; and Gabriele Kaiser, the Monograph Series Editor, for her support throughout the production process.

Cambridge, UK Andreas J. Stylianides
La Jolla, USA Guershon Harel

References

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

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

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

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

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

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

Stylianides, A. J., Bieda, K. N., & Morselli, F. (2016). Proof and argumentation in mathematics education research. In A. Gutiérrez, G. C. Leder, & P. Boero (Eds.), *The second handbook of research on the Psychology of Mathematics Education* (pp. 315–351). Rotterdam, The Netherlands: Sense Publishers.

Stylianides, G. J. (Ed.) (2014). Reasoning-and-proving in mathematics textbooks: From the elementary to the university level. *International Journal of Educational Research* [Special Issue], *64*, 63–148.

Stylianides, G. J., & Stylianides, A. J. (Eds.). (2017). Research-based interventions in the area of proof. *Educational Studies in Mathematics* [Special Issue], *96*(2), 119–274.

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

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

Contents

Part I
Epistemological Issues Related to Proof and Proving

Chapter 1
Reflections on Proof as Explanation

Gila Hanna

Abstract This chapter explores the connection between two distinct ways of defining mathematical explanation and thus of identifying explanatory proofs. The first is the one discussed in the philosophy of mathematics, in which a proof is considered explanatory when it helps account for a mathematical fact, clarifying why it follows from others. It is concerned with intra-mathematical factors, not with pedagogical considerations. The second definition is the one current among mathematics educators, who consider a proof to be explanatory when it helps convey mathematical insights to an audience in a manner that is pedagogically appropriate. This latter view brings cognitive factors very much into play. The two views of explanation are quite different. The chapter shows, however, citing examples, that insights from what are considered by philosophers of mathematics to be explanatory proofs can sometimes form a basis for explanatory proofs in the pedagogical sense and thus add value to the curriculum.

Keywords Mathematical proof · Mathematical explanation · Explanatory proof Proof teaching · Epistemology

Mathematical Explanations

Over the past four decades the philosophy of mathematics has shifted markedly away from a focus on the logical foundations of mathematics and towards a detailed study of mathematical practice. There has been a remarkable increase in publications on the concept of explanation in mathematics (with some ensuing influence on mathematics education as well). As pointed out by Mancosu (2011), the discussion of mathematical explanation has encompassed two areas:

G. Hanna (✉)
Ontario Institute for Studies in Education, University of Toronto, Toronto, Canada
e-mail: gila.hanna@utoronto.ca

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_1

> The first area addresses the problem of whether mathematics can play an explanatory role in the natural and social sciences. The second deals with the problem of whether mathematical explanations occur within mathematics itself. (para. 1)

By explanations within mathematics, often referred to as "intra-mathematical explanations", philosophers mean mathematical explanations of mathematical facts, as opposed to "scientific explanation using mathematics". This chapter will be concerned with intra-mathematical explanations and their implications for mathematics education.

When philosophers of mathematics discuss intra-mathematical explanations, they are unsurprisingly not concerned with pedagogy. They see such explanations simply as another facet of mathematics, one that is not designed with an audience in mind (other than mathematicians and philosophers of mathematics). They are correct in this view, because they use the term "explanation of a fact" in the sense of "accounting for a fact", similar to its use in the natural sciences. An example from the natural sciences might be an "explanation" of the occurrence of four seasons by pointing to the fact that the Earth tilts on its axis at an angle of 23.5° relative to our orbital plane, causing each hemisphere to be oriented toward the sun for half of the year and away from it for the other. The tilt of the Earth "explains" or "accounts for" the seasons. This type of explanation is meant as an essentially scientific explanation; it is not directed to a specific audience and it is not concerned with pedagogical considerations.

Intra-mathematical Explanations

The nature of mathematical explanation has long been a topic of debate among philosophers of mathematics (Cellucci 2008; Mancosu 2011; Steiner 1978). Though philosophers have come to a consensus on the central importance of mathematical explanation and accordingly have looked more closely at what may or may not count as explanation in mathematics, they widely diverge in their conclusions.

Some have gone as far as to argue that there are no explanations within mathematics. Zelcer (2013), for example, restates the accepted view that explanations in mathematics should be taken to be analogous to scientific explanations, "not mere stylistic features that communicate mathematics more clearly or in a psychologically more satisfying …way", but goes on to maintain that "given what we expect from a theory of explanation [meaning akin to scientific explanation], nothing comparable is possible in mathematics" (p. 176). This is a minority view. Most philosophers agree that intra-mathematical explanations do exist and think that most of them take the form of a proof.

If one were to take the position that an explanation is simply a deductive argument, then all proofs would automatically be explanations. However, almost all mathematicians make the very useful distinction between proofs that only demonstrate that a fact is true and proofs that also show why it is true. The latter are known as "proofs that explain". Certainly the distinction is not black and white, and indeed mathematicians do attribute to proofs various degrees of explanatory power (Mancosu 2011; Sandborg 1997).

In this chapter I take as a point of departure the distinction between proofs that merely prove and proofs that also explain, and I also accept the majority view that intra-mathematical explanations do exist, as evidenced by examples to be shown. (It is perhaps worth reiterating here that the term "proofs that explain" refers not to mere explanations, but to valid mathematical proofs that happen to possess the additional feature of explanatory power.)

Later I will survey three different philosophical accounts of intra-mathematical explanation, each of which offers criteria that can be used to identify mathematical explanations, and in particular explanatory proofs. These accounts deal with philosophical issues, not pedagogical ones, but this does not preclude them from having implications for the pedagogy of mathematics. In discussing each of these philosophical accounts, I will point out how some of the criteria they offer might be put to good use by mathematics educators in developing pedagogical explanations. To set the stage I will first discuss the nature of pedagogical explanation.

Pedagogical Explanations

Proof is a central feature of the mathematics curriculum, as it is of mathematics. Thus a key role of proving in the classroom is to teach proof itself: its use in justification, its strategies, its techniques, and its various forms. In this role, the most important goal of proof is to generate an understanding of the need to prove, of the process of proving, and of the role of deductive reasoning and logical inference. The focus of the present chapter, however, is on the potential complementary role of proofs in fostering a greater understanding of other mathematical concepts and propositions. As will be seen, it is often possible to find the happy concurrence in which a proof enlightens both the process of proving and the broader mathematical context with which it deals.

Where the classroom goal at hand is to generate an understanding of a mathematical proposition and the mathematical context in which it is embedded, and a proof is one vehicle to that end, the proof will clearly be most effective when it embodies explanation. The teaching of such a proof would be concerned not only with establishing its conclusion, but also with its main ideas, its overall structure, and its relationship to other mathematical fields and concepts (Balacheff 2010; De Villiers 2004; Hanna 1990, 2000; Mason and Hanna 2016).

It is no accident, then, that mathematics educators have been motivated to examine the idea of an explanatory proof. In this regard mathematics education does not differ from mathematics itself as much as might be thought. As mentioned, mathematicians too display concern for explanation. The mathematician Robinson (2000), for example, sees explanation as the most important role of proof. He describes mathematical proofs as having two potential components: proof-as-guaranty and proof-as-explanation. The former is concerned only with demonstrating the truth of a theorem, and is judged by its syntactical correctness. The latter component aims to shed light on why the theorem is true, and is judged by its explanatory value. As Robinson (2000) puts it, "The explanatory process caused by the cognitive internalization of a proof … seems to me to be both (far) more important, (far) more interesting, and (far) more challenging" (p. 280). His viewpoint is largely shared by mathematicians, so in this respect mathematics education is entirely reflective of mathematics itself.

In mathematical practice, of course, explanation is addressed by mathematicians to mathematicians, while in mathematics education it is addressed to a range of audiences. Understanding depends on the existing knowledge of the learners, on their developmental level, and on the quality of instruction (Tall et al. 2012; Harel and Sowder 2007; Mejia-Ramos et al. 2012). What it means to understand a proof may not admit of a tight definition, but it is generally agreed that it includes the ability to reproduce a proof, to identify its main idea(s), to see where in the proof certain assumptions are needed, to see why certain steps are essential, and to follow the deductive process (Avigad 2008; Hanna and De Villiers 2012; Stylianides and Stylianides 2009). Mathematics educators are also generally agreed that it is easier to meet these criteria of understanding when an explanatory proof can be enhanced by a visual representation (Clements 2014; Inglis and Mejía-Ramos 2009).

Research by Educators

As discussed, the focus of this chapter is on intra-mathematical explanation. It explores the insights of three philosophers of mathematics on the use of explanatory proofs to convey understanding among mathematicians, and evaluates these insights for their applicability to teaching. But this focus is not to imply that mathematics educators have not addressed the use of explanatory proofs directly from a pedagogical perspective.

In fact, educators have generated ample research literature on the pedagogical aspects of proof. Naturally, it has been largely motivated by issues that have arisen in the teaching and learning of proof at all levels, from elementary to tertiary. Many of the researchers have touched upon the complex relationship between proof and explanation. Their priority, however, has not been to offer specific criteria that might make a particular proof explanatory of a mathematical concept or proposition, but rather to suggest ways in which the process of proving itself might be made easier for students to grasp.

I will point out here only a few recent papers by educators that discuss this topic. Leron and Zaslavsky (2013), for example, proposed engaging students with the main ideas of a proof by presenting them with "generic proofs", while Selden and Selden (2015) argued for dividing the teaching of a proof into a "formal-rhetorical" part and a "problem-centered" one. Raman (2003) and Raman et al. (2009)

maintained that directing the attention of students to the "key ideas" of a proof could foster understanding. Hanna and Mason (2014) discussed the role of such key ideas in understanding a proof and in remembering and reconstructing it. Lastly, Stylianides et al. (2016) examined the conditions under which even a proof by mathematical induction can be explanatory.

On questions of proof and its explanatory nature there has already been a certain amount of collaboration between scholars in the philosophy of mathematical practice and those who focus on mathematics education (Hanna et al. 2010); it would be beneficial for both disciplines to continue the dialogue.

Philosophical Models of Explanation

In this section I will survey accounts by three philosophers of mathematics who examined intra-mathematical explanation with a view to identifying factors that make a proof explanatory to others in the discipline. The first two, Steiner (1978) and Kitcher (1981), are considered by other philosophers of mathematics to have offered most valuable insights into intra-mathematical explanation (Mancosu 2011), while the third, Lange (2014), has recently presented a new perspective on aspects of mathematical explanation. None of these three philosophers were concerned with pedagogical implications. In the last section, however, I will provide some examples that show how factors they identified can nevertheless provide guidance to mathematics educators in choosing explanatory proofs for use in the curriculum.

Mark Steiner: A Characteristic Property

Steiner (1978) presented a model of mathematical explanation that draws upon the distinction between explanatory and non-explanatory proofs. Seeking to identify the factors that make a proof explanatory, he first examined Feferman's attribution of explanatory power to abstractness and generality. Feferman had stated that "Abstraction and generalization are constantly pursued as the means to reach really satisfactory explanations which account for scattered individual results" (Feferman quoted in Steiner 1978, p. 135), and more concisely that "Of two proofs of the same theorem, the *more* explanatory is the more abstract (or general)" (p. 136). Steiner showed through a few examples, however, that Feferman's criteria fail to account for explanatory power (p. 143).

In Steiner's view, a proof is explanatory when it reveals and makes use of the mathematical ideas that motivate it, that is, when it makes evident that a "characterizing property" is responsible for making the conclusion true. The "characterizing property" will change from theorem to theorem, because it is unique to a given mathematical entity.

As Steiner put it, “an explanatory proof makes reference to a characterizing property of an entity or structure mentioned in the theorem, such that from the proof it is evident that the result depends on the property” (1978, p. 143). This criterion is in effect an answer to the question: “What is it about the proof that makes it possible to see how the conclusion follows naturally?”

As an example of a proof that makes use of a characterizing property, Steiner presents the proof of the equation $S_n = n(n + 1)/2$ commonly known as the Gaussian proof. In this case the characterizing property is that of symmetry. The proof proceeds by adding the sequence “$1 + 2 + \cdots + n$” to the same sequence reversed, “$n + (n - 1) + \cdots + 1$”, to give “$(n + 1) + (n + 1) + \cdots + (n + 1) = 2S_n = n(n + 1)$”. This proof, using symmetry, shows why the result is true, and is thus “more illuminating” than a proof by mathematical induction.

Steiner goes on to show that a geometrical proof accompanied by a visual demonstration of the same sum is an “even more explanatory proof” (p. 137). He is ambivalent, however, about asserting a more general link between explanatory power and the ability to visualize a proof. While admitting that many explanatory proofs do rely on some “pictorial aspects” such as diagrams, he states that this criterion “is too subjective to excite” (p. 143).

Steiner’s model was endorsed by several philosophers of mathematics, among them Weber and Verhoeven (2002) who went on to offer a refined and improved version. On the other hand, a number of others in the field found much to criticize in Steiner’s model. Resnik and Kushner (1987) provided a counter-example and argued that the choice of a characterizing property is bound to be arbitrary. Hafner and Mancosu (2005) showed that in the case of the proof of Kummer’s convergence criterion, the explanatory power of the proof cannot be accounted for by Steiner’s criteria. Molinini (2012) examined the theorem known as the “Euler theorem” that appears in “Découverte d’un nouveau principe de mécanique” (1750) and discussed it as it relates to the notion of explanatory proof in mathematical practice. He then argued that the criteria for mathematical explanation proposed by Steiner are insufficient. This is what Molinini (2012, p. 123) had to say about Steiner’s model:

> … it might be more philosophically profitable to abandon Steiner’s idea that an explanatory proof depends on a particular property of an entity mentioned in the theorem in favour of an approach which focuses on the preferences expressed by the mathematicians for some mathematical concepts or for the particular mathematical framework used to prove a theorem. On the other hand, it might be thought that the notion of explanatory proof cannot be captured *simpliciter,* as Steiner proposes, but that there is a variety of explanatory proof-practices in mathematics.

Although Steiner and the others who assessed his model did not concern themselves with pedagogical explanation, I will show later, as mentioned, that aspects of his model of mathematical explanation are quite relevant to mathematics education.

Philip Kitcher: Theoretical Unification

Kitcher (1981, 1989) approaches the concept of explanation by stating that:

> …successful explanations … belong to a set of explanations, *the explanatory store*, …. Intuitively the explanatory store associated with science at a particular time contains those derivations which collectively provide the best systematization of our beliefs. (Kitcher 1989, p. 430)

He does not use the term "proofs that explain", but he does assert that the ideal explanations are derivations (though of course not all derivations are explanations).

For Kitcher (1981) a derivation is an acceptable explanation for its conclusion when it is capable of systematizing a set of statements. This means that the derivation identifies a set of patterns that can be used repeatedly, "… and, in demonstrating this, it teaches us how to reduce the number of facts we have to accept as ultimate (or brute)" (p. 432). This leads to what Kitcher refers to as unification, which he regards as a form of explanation.

According to this approach, a mathematical fact (or scientific fact, for that matter) is "explained" by showing that it is part of a larger set of facts that share common patterns, which he calls "unifying explanatory patterns". Thus "… to explain is to fit the phenomena into a unified picture insofar as we can" (Kitcher 1981, p. 500). This view of intra-mathematical explanation will reverberate with mathematics educators who are pleased if they are able to illuminate a theorem and its proof by reference to related bodies of mathematical knowledge shared by their students.

Kitcher's model may not be as relevant to mathematics education as it might appear, however, because it is difficult to assign to unification per se—fitting a mathematical fact into a unified picture—any degree of explanatory power in the cognitive sense. It seems, too, that the methods that might lead to a potential unification are not necessarily ones that are likely to offer an explanation of the mathematical fact. The difference between the explanatory process and the process of unification was put by Halonen and Hintikka (1999, pp. 27–28) as follows:

> The explanatory process is geared to a particular explanandum and a particular background theory. The crux of any one explanatory process lies in finding the ad hoc premises from which (together with the background theory) the particular explanandum in question follows. There is not place for any unification in such a process. … In other words, it is not a matter of explanation, in the sense that it should affect our ideas about the process of explanation. Instead it is a matter of theory formation and theory selection.

Marc Lange: Symmetry, Unity, and Salience

Lange (2014), embracing the distinction between proofs that only prove and proofs that both prove and explain, investigated what makes a proof explanatory. He sees the process of proof as key:

> … I will argue that two mathematical proofs may prove the same theorem from the same axioms, though only one of these proofs is explanatory. My goal in this essay will be to identify the ground of this distinction. Accordingly, my focus will be on the course that a given proof takes between its premises and its conclusion. The distinction between explanatory and nonexplanatory proofs from the same premises must rest on differences in the way they extract the theorem from the axioms. (p. 487)

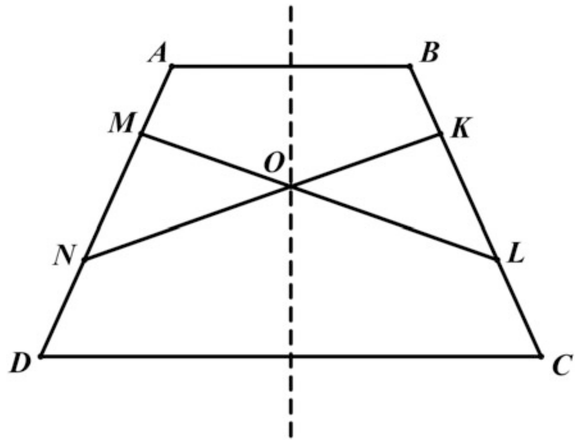
Lange goes on to argue that a proof can be explanatory only if "some feature of the result is salient" and the proof builds upon that salient feature (Lange 2014, p. 489). According to him "salience" is a feature that is "worthy of attention" and that "helps to determine what a proof must do in order to explain why the theorem holds" (p. 488). When the salient feature of the result is symmetry, for example, "a proof is privileged as explanatory because it exploits a symmetry in the problem—a symmetry of the same kind as initially struck us in the fact being explained" (p. 499). However, Lange does not restrict a salient feature to symmetry alone; he admits that some feature "… other than its symmetry could likewise be salient, prompting a why question answerable by a proof deriving the result from a similar feature of the given" (p. 507).

Lange gives several examples of this situation from number theory and geometry. For example, as shown in Fig. 1.1, symmetry is the salient feature of the theorem: If ABCD is an isosceles trapezoid (AB parallel to CD, AD = BC) such that AM = BK and ND = LC, then ML = KN. Thus of all the proofs of this theorem, the one that Lange considers explanatory is the one that makes use of a symmetry argument (Fig. 1.1). Lange (2014) adds:

> The theorem (that ML = KN) "makes sense" in view of the figure's overall symmetry. Intuitively, a proof that fails to proceed from the figure's symmetry strikes us as failing to focus on "what is really going on": that we have here the same figure twice, once on each side of the line of symmetry. (p. 502)

His conclusion is that one can speak of a proof that explains only "in a context where some feature of the result being proved is salient" (p. 507). He allows for

Fig. 1.1 A proof that is explanatory because it exploits an overall symmetry



salient features other than symmetry, but maintains that one cannot expect to find an explanatory proof for a result that has no salient feature. The examples discussed below illustrate the value of a salient feature, but not all mathematicians or educators would agree with Lange that such a feature is the only road to explanatory power.

The role of symmetry in mathematical proofs, advocated by both Steiner and Lange, is also strongly supported by Giuseppe Longo, a philosopher of mathematics with an interest in the cognitive foundations of mathematics. In his review of epistemological perspectives on mathematical concepts and proofs, Longo (2011, p. 64) affirmed its importance:

> I have shown how geometric judgements penetrate proof even in number theory; I argue, a fortiori, their relevance for general mathematical proofs. We need to ground mathematical proofs also on geometric judgments which are no less solid than logical ones: "Symmetry", for example, is at least as fundamental as the logical "modus ponens"; it features heavily in mathematical constructions and proofs. (p. 64)

Proofs Considered Pedagogically Explanatory

The following are proofs that show *why* a result is true in a manner that would be entirely suitable for classroom use, even though the keys to their explanatory power stem from the reflections of the philosophers of mathematics discussed above—who were addressing intra-mathematical explanation in the absence of pedagogical considerations.

Example 1: Pick's Theorem

Pick's theorem provides a simple formula for calculating the area A of a polygon in terms of the number i of *lattice points in the interior* of the polygon and the number b of lattice points on the boundary: $A = i + b/2 - 1$.

The theorem is named after Georg Pick, who published it in 1899. It is easy to state and verify, but it is not obviously true. It has been proved in different ways several times. The reason for reproving is usually to create a proof that is more explanatory, more elegant, more rigorous, or more general (Sandborg 1997). Some of these proofs use mathematical induction, some use dissections into minimal triangles followed by an application of Euler's polyhedron formula, while others use partitioning and arguments from graph theory. Blatter (1997) offered yet another proof, one of a conceptional nature, in the form of a thought experiment (*Gedankenexperiment*).

Although the theorem is easily understood, most of its proofs were not within the grasp of high school students. The following explanatory proof was provided by middle- and high-school students at St. Mark's Institute of mathematics (Tanton 2010) who sought answers to the following questions: (1) "Why are interior points each worth 1?", (2) "Why are boundary points each worth ½?", and (3) "Why is there a −1?" (p. 34).

The explanatory proof developed by the students could be seen as an application of Lange's views on intra-mathematical explanation. As discussed below, the students took notice of "salient features" of the theorem and were prompted to formulate three appropriate "why" questions "answerable by a proof deriving the result from a similar feature of the given" (Lange 2014, p. 507).

The students started by examining a lattice rectangle as shown in Fig. 1.2. The following is a summary of their reasoning:

> Each interior point contributes one full square unit of area and each boundary point different from a vertex half a unit of area. If we extend the sides of the rectangle to make its exterior angles explicit we can introduce additional area so that each vertex also contributes half a unit of area. As the exterior angles of any polygon sum to one full turn, this excess in area amounts to one full square unit. The "−1" in Pick's formula compensates for this. (Tanton 2010, p. 34)

They then applied the same reasoning to any simple lattice polygon (Fig. 1.3), pointing out that:

> …diagonal line segments connecting two lattice points are rotationally symmetric about their midpoints. In particular, any cell (unit square) that is intercepted by such a diagonal and divided into two parts is matched by a rotationally symmetric cell divided into the same two parts. (And the matching portions are on alternate sides of the diagonal) (p. 34)

From here they reached the proof of Pick's area theorem for any polygon, one which they considered to be a proof without words (Fig. 1.4). They then proceeded to generalize the results.

This proof, designed for classroom use, nicely demonstrates the explanatory criterion of symmetry suggested by the philosophers of mathematics, Lange, Longo, and Steiner, as well as the criterion of salience suggested by Lange.

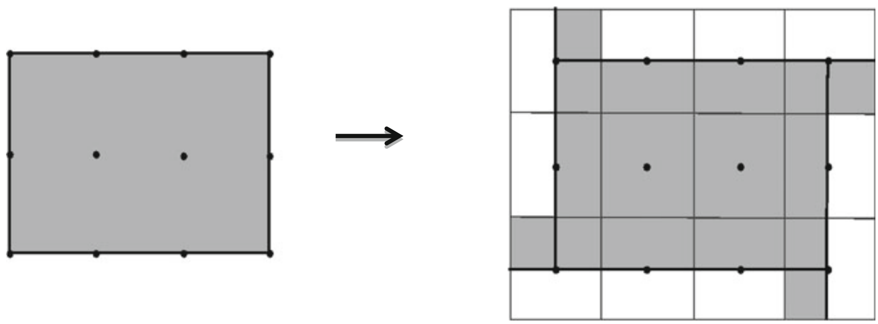


Fig. 1.2 Lattice rectangle where each boundary point is surrounded with a ½ unit square and showing the excess area which amounts to one full square unit (thus −1)

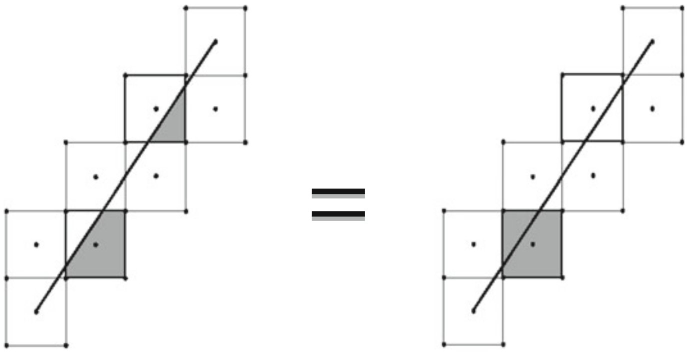


Fig. 1.3 Each portion outside a unit square with center in the interior of the polygon is replaced by an equal portion through a rotational symmetry about the diagonal midpoint

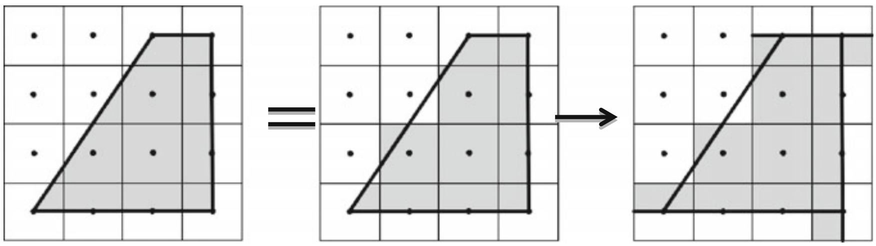


Fig. 1.4 Explanatory proof of $A = i + b/2 - 1$

Example 2: The Irrationality of $\sqrt{2}$

Here are two indirect proofs (that is, proofs by contradiction: proving *p* by showing that *not p* leads to a contradiction). Most would consider indirect proofs as inherently not explanatory, but this is not necessarily the case. Although the literature of mathematics education seems to show that many students find indirect proofs unenlightening and somewhat cognitively demanding (Antonini and Mariotti 2008), it is possible to present some such proofs in ways that show *why* a result is true.

The following two indirect proofs do show how proving the result (that $\sqrt{2}$ is not rational) can be made explanatory by revealing the heart of the matter, in particular when there is a visual aid. The two proofs are both intra-mathematically and pedagogically explanatory.

2a. Carpet proof of the irrationality of $\sqrt{2}$

There are over 30 different proofs of the irrationality of $\sqrt{2}$. This geometric proof, often referred to as the "Carpet proof", is attributed to Tennenbaum (Conway and

Shipman 2013; Miller and Montague 2012). It is a good illustration of how "geometric judgements penetrate proof even in number theory" (Longo 2011, p. 64).

Assume $\sqrt{2}$ is rational. Select the smallest integer-sided square m^2 whose area is the sum of two integer sided congruent squares, so that $m^2 = 2n^2$, which means that an "m $\times$ m" square has the same area as two "n $\times$ n" ones. As shown in Fig. 1.5, place two smaller square carpets, of size n^2 each, in opposite corners of the larger square carpet of size m^2.

Then the area of overlap will be a square equal in area to the total area not covered by the two carpets, which is made up of two congruent squares. (This is so because the statement says that the area of square m^2 is equal to the area of $2n^2$.) The side of each uncovered square is $(m - n)$ while the side of the overlap is $(2n - m)$. Clearly, then, the sides of these two uncovered congruent squares and the overlap square are also whole numbers smaller than the original ones, thus contradicting the assumption of a smallest presentation of $\sqrt{2}$ as a rational number.

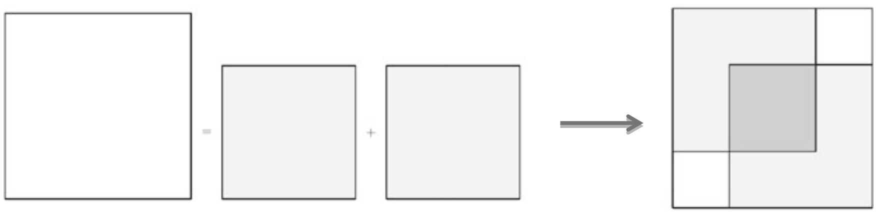


Fig. 1.5 Carpet proof of the irrationality of $\sqrt{2}$

2b. Proof of the irrationality of $\sqrt{2}$ based on parity

The following school proof is based on the concept of parity. This salient feature of the proof (Lange 2014) happens to be an elementary concept familiar to students. The explanatory power of this particular proof, also known as the Pythagorean proof, has been discussed by Steiner (1978) who showed that it is explanatory in the eyes of philosophers of mathematics. It also happens to be a valuable explanatory proof when viewed with pedagogical considerations in mind.

Suppose that $\sqrt{2}$ were rational. Then $\sqrt{2} = m/n$ for some integers m, n in lowest terms, i.e., m and n have no common factors. Then $2 = m^2/n^2$, which implies that $m^2 = 2n^2$. Hence m^2 is even, which implies that m is even. Then $m = 2k$ for some integer k. So $2 = (2k)^2/n^2$, but then $2n^2 = 4k^2$, or $n^2 = 2k^2$. So n^2 is even. But this means that n must also be even, because the square of an odd number cannot be even. We have just shown that both m and n are even, which contradicts the fact that m, n are in lowest terms. Thus our original assumption (that $\sqrt{2}$ is rational) is false, so $\sqrt{2}$ must be irrational.

Example 3: The Product of Any Three Consecutive Nonzero Natural Numbers Is Divisible by 6

This result appears in most high-school mathematics curricula. It is often proved by mathematical induction. The intra-mathematical explanatory proof provided by Lange (2014, p. 510), however, exploits a salient feature "common to every triple of consecutive nonzero natural numbers" and the 2-line proof is immediate: "Of any three consecutive nonzero natural numbers, at least one is even (that is, divisible by 2) and exactly one is divisible by 3. Therefore, their product is divisible by $3 \times 2 = 6$."

In this case the criterion of salience makes the proof explanatory not only intra-mathematically but also pedagogically. It clearly shows *why* the claim is true in a way that would be judged by most educators to potentially lead to a good understanding of the proof. An interesting discussion of two proofs of the more general theorem that states: "The product of any k consecutive positive integers $n(n+1)\ldots(n+k-1)$ is divisible by $k!$" can be found in a blog post by Gowers (2010).

Conclusion

The chapter has shown that in the practice of mathematics there is such a thing as intra-mathematical explanation (which ignores issues of pedagogy), and has then discussed several ideas put forward by philosophers of mathematics as to what features of a proof would make it "intra-mathematically explanatory". Though these features were not identified and explored for their pedagogical value and thus would not have necessarily come to the attention of mathematics educators, the chapter has gone on to show that many of them can nevertheless be of help to mathematics educators in identifying or creating proofs that are explanatory in the pedagogical sense.

Three current and influential models of explanation have been discussed, but of course other valuable models may yet be formulated. Indeed, the notion of intra-mathematical explanation in general, and in particular the distinction between a proof that shows that a result is true and a proof that also explains *why* a result is true, have been topics of intensive debate among philosophers, in particular among philosophers of mathematical practice. Unfortunately, to quote Burgess, this has shown "… rather meager and inconclusive results… And needless to say there are many quotable things mathematicians have said about such a distinction at one time or another too, not all by any means pointing in the same direction." (Burgess 2014, p. 1347). For example, in Cellucci's opinion heuristics and the important connection between explanation and discovery have so far been overlooked:

> What is crucial in a mathematical explanation is not a characterizing property of an entity or structure mentioned in the theorem, but rather the heuristic value of the hypothesis, its effectiveness as a means of discovery. While characterizing properties are properties that entities or structures mentioned in the problem are supposed to possess, the heuristic value of a hypothesis may depend on entities or structures not mentioned in the problem. (Celluci 2008, p. 207)

It seems clear that the notions of mathematical explanation and of explanatory proof cannot be captured by a single model. Potential additional models for explanation in mathematics would most certainly take into consideration approaches such as the use of analogies, examples, and rule-based logic, as well as assess the roles of intuition and visualization, perhaps while entertaining varying criteria for the validity of a proof. Educators will have to continue to draw on multiple resources, rely on their judgment, and be pragmatic when seeking to identify or construct proofs that are sufficiently explanatory to meet their pedagogical goals.

Acknowledgements Preparation of this chapter was supported in part by the Social Sciences and Humanities Research Council of Canada.

References

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

Avigad, J. (2008). Understanding proofs. In P. Mancosu (Ed.), *The philosophy of mathematical practice* (pp. 317–353). Oxford, UK: Oxford University Press.

Balacheff, N. (2010). Bridging knowing and proving in mathematics: A didactical perspective. In G. Hanna, H. N. Jahnke, & H. Pulte (Eds.), *Explanation and proof in mathematics* (pp. 115–135). New York, NY: Springer.

Blatter, C. (1997). Another proof of pick's area theorem. *Mathematics Magazine, 70*(3), 200.

Burgess, J. P. (2014). Review of the book *Why is there philosophy of mathematics at all?* by I. Hacking. *Notices of the AMS, 61*(11), 1345–1348.

Cellucci, C. (2008). The nature of mathematical explanation. *Studies in History and Philosophy of Science Part A, 39*(2), 202–210.

Clements, M. K. A. (2014). Fifty years of thinking about visualization and visualizing in mathematics education: A historical overview. In M. Fried & T. Dreyfus (Eds.), *Mathematics & mathematics education: Searching for common ground* (pp. 177–192). Netherlands: Springer.

Conway, J., & Shipman, J. (2013). Extreme proofs I: The irrationality of $\sqrt{2}$. *The Mathematical Intelligencer, 35*(3), 2–7.

De Villiers, M. (2004). Using dynamic geometry to expand mathematics teachers' understanding of proof. *International Journal of Mathematical Education in Science and Technology, 35*(5), 703–724.

Gowers, T. (2010). *Gowers's weblog: Mathematics related discussions. Are these the same proofs?* https://gowers.wordpress.com/2010/09/18/are-these-the-same-proof/.

Hafner, J., & Mancosu, P. (2005). The varieties of mathematical explanation. In P. Mancosu, K. Jørgensen, & S. Pedersen (Eds.), *Visualization, explanation and reasoning styles in mathematics* (pp. 215–249). Berlin: Springer.

Halonen, I., & Hintikka, J. (1999). Unification – it's magnificent but is it explanation? *Synthese, 120*(1), 27–47.

Hanna, G. (1990). Some pedagogical aspects of proof. *Interchange, 21*(1), 6–13.
Hanna, G. (2000). Proof, explanation and exploration: An overview. *Educational Studies in Mathematics, 44*(1–2), 5–23.
Hanna, G., & De Villiers, M. (Eds.). (2012). *Proof and proving in mathematics education*. New York, NY: Springer.
Hanna, G., Jahnke, H. N., & Pulte, H. (Eds.). (2010). *Explanation and proof in mathematics: Philosophical and educational perspectives*. New York: Springer.
Hanna, G., & Mason, J. (2014). Key ideas and memorability in proof. *For the Learning of Mathematics*, 12–16.
Harel, G., & Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning, 2* (pp. 805–842). Charlotte, NC: Information Age Publishing.
Inglis, M., & Mejía-Ramos, J. P. (2009). On the persuasiveness of visual arguments in mathematics. *Foundations of Science, 14*(1–2), 97–110.
Kitcher, P. (1981). Explanatory unification. *Philosophy of Science, 48,* 507–531.
Kitcher, P. (1989). Explanatory unification and the causal structure of the world. In P. Kitcher & W. Salmon, (Eds.), *Scientific explanation* (pp. 410–505). (Minnesota Studies in the Philosophy of Science, Vol. XIII), Minneapolis: University of Minnesota Press.
Leron, U., & Zaslavsky, O. (2013). Generic proving: Reflections on scope and method. *For the Learning of Mathematics, 33*(3), 24–30.
Longo, G. (2011). Theorems as constructive visions. In G. Hanna & M. deVilliers, (Eds.), *Proof and proving in mathematics education* (pp. 51–66). New York, NY, Springer.
Lange, M. (2014). Aspects of mathematical explanation: Symmetry, unity, and salience. *Philosophical Review, 123*(4), 485–531.
Mancosu, P. (2011). Explanation in mathematics. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy.* http://plato.stanford.edu/entries/mathematics-explanation/.
Mason, J., & Hanna, G. (2016). Values in caring for proof. In B. Larvor (Ed.), *Mathematical cultures* (pp. 235–257). Berlin: Springer International Publishing.
Mejia-Ramos, J. P., Fuller, E., Weber, K., Rhoads, K., & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics, 79*(1), 3–18.
Miller, S., & Montague, D. (2012). Picturing Irrationality. *Mathematics Magazine, 85*(2), 110–114.
Molinini, D. (2012). Learning from Euler. From mathematical practice to mathematical explanation. *Philosophia Scientiae, 16*(1), 105–127.
Raman, M. (2003). Key ideas: What are they and how can they help us understand how people view proof? *Educational Studies in Mathematics, 52,* 319–325.
Raman, M., Sandefur, J., Birky, G., Campbell, C., & Somer, K. (2009). "Is that a proof?" Using video to teach and learn how to prove at the university level. In F.-L. Lin, F.-J. Hsieh, G. Hanna & M. De Villiers (Eds.), *Proof and proving in mathematics education: ICMI Study 19 Conference Proceedings, Taipei. Taiwan* (Vol. 2, pp. 154–159).
Resnik, M. D., & Kushner, D. (1987). Explanation, independence and realism in mathematics. *The British Journal for the Philosophy of Science, 38*(2), 141–158.
Robinson, J. A. (2000). Proof = guarantee + explanation. In S. Hölldobler (Ed.), *Intellectics and computational logic* (pp. 277–294). Dordrecht, The Netherlands: Kluwer.
Sandborg, D. (1997). *Explanation and mathematical practice*. Unpublished doctoral dissertation, University of Pittsburgh.
Selden, J., & Selden, A. (2015). A theoretical perspective for proof construction. In K. Krainer & N. da Vondrov (Eds.), *CERME 9—Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education, Prague, Czech Republic* (pp. 198–204).
Steiner, M. (1978). Mathematical explanation. *Philosophical Studies, 34,* 135–151.
Stylianides, G. J., Sandefur, J., & Watson, A. (2016). Conditions for proving by mathematical induction to be explanatory. *The Journal of Mathematical Behavior, 43,* 20–34.

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

Tall, D., Yevdokimov, O., Koichu, B., Whiteley, W., Kondratieva, M., & Cheng, Y. H. (2012). Cognitive development of proof. In G. Hanna & M. deVilliers (Eds.), *Proof and proving in mathematics education* (pp. 13–49). New York, NY: Springer.

Tanton, J., & With St. Mark's Students. (2010). Pick's Theorem—and beyond! *MAA Focus, 30*(1), 34–35. See also http://www.jamestanton.com/wp-content/uploads/2009/04/picks_theorem_focus_web-version.pdf.

Weber, E., & Verhoeven, L. (2002). Explanatory proofs in mathematics. *Logique et Analyse, 45,* 299–307.

Zelcer, M. (2013). Against mathematical explanation. *Journal for General Philosophy of Science, 44*(1), 173–192.

Chapter 2
Working on Proofs as Contributing to Conceptualization—The Case of IR Completeness

Viviane Durand-Guerrier and Denis Tanguay

Abstract In this chapter, we propose a mathematical and epistemological study about two classical constructions of the real number system, by Dedekind (cuts) and Cantor (Cauchy sequences), and the associated proofs of its completeness. In addition, we present two contrasting constructions leaning on decimal expansions. Our analysis points out that Dedekind's construction fosters a conceptualization of the real numbers leaning strongly on the total ordering of $\mathbb{Q}$ and $\mathbb{R}$, while putting aside the metrical aspects. By contrast, the more intricate construction through Cauchy sequences calls on complex objects, but yields to a better understanding of the topological relationship between rational and real numbers. We argue that suitable considerations of decimal expansions and of approximation issues enable to connect and complement those two approaches. These analyses highlight the dialectical interplay between syntax and semantics and the crucial role of the definitions of objects at play in proof and proving. The general didactical issue pertains to the potential contribution of analyzing proofs as a means for deepening the understanding of the related objects and of their ensuing conceptualization. We hypothesize that doing so with Dedekind's cuts, Cauchy sequences and decimal expansions open paths towards improving the conceptualization of the real numbers, by taking into account the triad discreteness/density/continuity.

Keywords Real numbers · Density · Continuity · Completeness
Cuts · Cauchy sequences · Syntax and semantics · Proof and conceptualization

V. Durand-Guerrier (✉)
Institut Montpelliérain Alexander Grothendieck, CNRS,
University of Montpellier, Montpellier, France
e-mail: viviane.durand-guerrier@umontpellier.fr

D. Tanguay
Université du Québec à Montréal, Montreal, Canada

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_2

Introduction

In this chapter, we present the first elements of a larger research project about the relationship between the choices made in constructing and defining mathematical objects, and the possible proofs induced by these choices. This project is based on the hypothesis that working with proof is likely to contribute to conceptualization by prompting a work with the mathematical objects at stake, in agreement with the syntax-semantics dialectic in proof and proving (e.g. Alcock and Weber 2004). To initiate this research project, we decided to focus on the concept of continuity, as it is conveyed by completeness of the real number system. This choice is motivated by epistemological and didactical considerations.

Ancient Greeks were already aware of the existence of incommensurable magnitudes whose ratios cannot be expressed as quotients of integers, but such ratios were not designated as *numbers*, even though methods—such as Eudoxus' *equimultiples*—were developed to compare and process them. With the birth of Calculus in the seventeenth century, the recourse to real numbers grew insofar as being systematized in mathematics, well before the emergence of explicit constructions, these dating back to the second half of the nineteenth century with, among others, the ones from Dedekind and Cantor. The need for these constructions became felt with the need for stronger theoretical bases supporting theorems such as the *Intermediate Value Theorem*. The latter was largely used, but its justification only leaned on graphical evidences, or (implicitly) drew on intuitive results such as the existence of a limiting value for any "magnitude [that] grows continually but not beyond all limits" (Dedekind 1963a, p. 24). It is the existence of such limiting values that Dedekind planned to prove "in a purely arithmetical fashion", as he will qualify his own approach. These existential statements, utilized as 'in-acts-theorems' (in Vergnaud's sense 1990), are indeed false in the set $\mathbb{Q}$ of rational numbers, which is incomplete in that respect. For Cantor and Dedekind, the explicit stake was to extend $\mathbb{Q}$ so to be able to prove analytically these results without resorting to geometrical evidences, and this amounts to ascertain completeness (Benis Sinaceur 2008b, pp. 45–46). Other constructions of $\mathbb{R}$ will come out afterwards, for instance as the set of infinite decimal expansions, as the power set of $\mathbb{N}$, or as the set of paths in an infinite binary tree (to be related to binary expansions). Finally, during the twentieth century, the real number system has been characterized by its structure of continuous totally ordered field.

In classrooms, teaching approaches vary depending on countries, schooling levels or even mathematical topics. As far as France is concerned, while during the "New Math" period, a construction of $\mathbb{R}$ through decimal expansions was proposed at the secondary level (e.g. Lelong-Ferrand 1964), no construction of $\mathbb{R}$ has been kept at this level by the reform that followed, partly in reaction to the New Math movement which was assessed as badly fitted to teaching and learning. At the *Université de Montpellier*, for several years, the real number system is introduced in an axiomatic form, the Supremum Axiom[1] having been selected to characterize completeness. The definition through infinite decimal expansions is sometimes proposed, in parallel with specifying the nature of real numbers.

We hypothesize that whatever the adopted approach, the notion of completeness remains the most difficult to conceptualize. Bergé (2010), for one, stresses the difficulties met by undergraduate students having taken four courses in Analysis and who replied to a questionnaire.

> For most of the students, doing typical exercises involving the supremum does not lead to the understanding that $\mathbb{R}$ is the set that contains all the suprema of its bounded above subsets. Few students can perceive that the notion of Cauchy sequences comes from the necessity of characterizing the kind of sequences that 'must' converge—an essential insight required to further develop mathematical analysis—and that completeness is related to the issue whether a limit is guaranteed to lie in $\mathbb{R}$. (Op. cit., p. 226)

In our work, we assume that working on proof is fruitful as regards the conceptualization processes in mathematics (e.g. Frege 1971). This assumption motivates the trend of our research, which falls within a larger project in mathematical didactics about the conceptualization of real numbers, and the consideration of the triad discreteness/density/continuity, instead of the mere dyad discreteness/continuity. The general didactical issue under study is: what in-depth work on the objects at stake can be envisaged through analyzing proofs?

To carry out such a research, it is necessary to take into account dialectical interplays between mathematical, epistemological and didactical studies. We propose here a first reflection work where mathematical and epistemological considerations are networked, as prolegomena to the didactical studies. We will look into the two constructions of $\mathbb{R}$ having been historically the first formally set up, and will examine to what proof of the Supremum Theorem (or Axiom) these constructions lead to, according to the authors under consideration. In addition, we will present two contrasting constructions relying on decimal expansions.

Definition of Real Numbers by Cuts

Among the first 'constructions' of $\mathbb{R}$ to appear during nearly the same period around 1870 (Meray, Cantor, Heine, Kossac and Weierstrass…), the one by Dedekind (1872) is most likely the more formally accomplished. Let us recall some of its salient features.

[1]Any nonempty subset of $\mathbb{R}$ which is bounded above has a supremum (or least upper bound).

Creation of Irrational Numbers

A cut is a partition (A_1, A_2) of $\mathbb{Q}$ such that for all $a_1 \in A_1$, $a_2 \in A_2$, we have $a_1 < a_2$. Any rational q induces two cuts, considered as 'non-essentially-distinct', one in which q is the maximum of A_1 and one in which q is the minimum of A_2. The 'non-completeness' of $\mathbb{Q}$ is due to the fact that there exist cuts that are not induced by rational numbers, for instance the cut in which $A_2 = \{q \in \mathbb{Q} | 0 < q \text{ and } 2 < q^2\}$ and A_1 is its complement (in $\mathbb{Q}$). Dedekind shows that any integer that is not the square of an integer gives rise to such a cut, these cuts (bound to correspond to irrational numbers) thus being infinitely many. Indeed, for every cut that is not produced by a rational number, Dedekind creates a new *irrational* number, unequivocally defined by this cut. He then states that to each cut corresponds a *real number*, rational or irrational, and from this point onwards he designates by $\mathbb{R}$ the 'system' of all these numbers (Dedekind 1963a, pp. 15–19). Let us look into the way Dedekind 'recovers' ordering, and extends it to this new set by resorting decisively to the density[2] (for the usual order) of $\mathbb{Q}$. Let (A_1, A_2) and (B_1, B_2) be two cuts. If A_1 and B_1 are not equal (as sets), then there exists $\alpha_1 \in A_1$ which is not in B_1. Then $\alpha_1 = \beta_2$ for a $\beta_2 \in B_2$, since (B_1, B_2) is a partition. If α_1 is the only member of A_1 in this case, then it is easy to show that α_1 is the maximum of A_1 and β_2 the minimum of B_2, so that the cuts (A_1, A_2) and (B_1, B_2) are non-essentially-distinct, and equal in $\mathbb{R}$. If there exist at least two elements α_1 and α'_1 in A_1 which are not in B_1, then the infinity of rational numbers between α_1 and α'_1 (density of $\mathbb{Q}$) are at the same time in A_1 and in B_2. In that case, Dedekind says that "the numbers α and β corresponding to these two essentially different cuts (A_1, A_2) and (B_1, B_2) are *different*, and further that α is *greater* than β, that β is *less* than α, which we express in symbols by $\alpha > \beta$ as well as $\beta < \alpha$" (Dedekind 1872, 1963a, b, p. 17, emphasis in original). Dedekind shows that it is a total order on $\mathbb{R}$. Clearly, it corresponds to the usual order on $\mathbb{Q}$ when the numbers induced by the cuts are rational.

Continuity of the Set of Real Numbers

Dedekind shows afterwards that the set $\mathbb{R}$ thus obtained is 'continuous'—we would rather say 'complete' nowadays—by showing that any cut of $\mathbb{R}$ is induced by a unique $\rho \in \mathbb{R}$. We may assess the elegance of the construction in the light of the simplicity of this proof, which goes as follows. Suppose that (A_1, A_2) is a cut of $\mathbb{R}$. It gives rise to a cut $(A_1 \cap \mathbb{Q}, A_2 \cap \mathbb{Q})$ of $\mathbb{Q}$, which is denoted A'_1, A'_2, and to it there corresponds an element of $\mathbb{R}$ denoted ρ. If σ is any number distinct from ρ, then we

[2]We are referring here to *density-in-itself*, or intrinsic density with respect to order <, according to which between two rational numbers there is always a third one (different from the first two) and hence, infinitely many.

consider $q \in \mathbb{Q}$ among the infinity of rationals between ρ and σ (cf. the discussion on order). If $\sigma < q < \rho$, then q is in A'_1 and thus also in A_1, so much that σ is also in A_1. If $\rho < q < \sigma$, it is the contrary and $\sigma \in A_2$. So any element σ of $\mathbb{R}$ is either smaller than ρ and in A_1 or greater than ρ and in A_2, and ρ is indeed the unique real number performing the cut (A_1, A_2).

Dedekind then shows that the operations + and × on $\mathbb{Q}$ can adequately be extended to $\mathbb{R}$ and are compatible with the order on $\mathbb{R}$: arguments that are essentially based on order and 'continuity' (in Dedekind's sense), and that are in fact undertaken formally only for addition.

Proofs of the Main Theorems Leaning on Cuts and Completeness

Dedekind comes finally to the very motivation of the construction, which is to prove the standard theorems of 'infinitesimal analysis' with the sole recourse to the 'arithmetic' (as he states it himself) of the new set, constructed as an extension of $\mathbb{Q}$. In this way, he will show first "one of the most important theorem [that] may be stated in the following manner: 'if a magnitude x grows continually but not beyond all limits, it approaches a limiting value'" (Dedekind 1963a, pp. 24–25). In the current terminology, we would say that any bounded above increasing (real) function f has a limit. The proof consists in considering, on the one hand the set M of upper bounds for the images of f, on the other hand the set M' of numbers that are not upper bounds, and to show that (M', M) is a cut of $\mathbb{R}$. By the just established continuity of $\mathbb{R}$, this cut is performed by a unique real number, and it is shown that this number is the expected limiting value. With similar arguments although somewhat more intricate, Dedekind shows afterwards that a function, satisfying the analog of the 'Cauchy criterion' for sequences, has a limit (in $\mathbb{R}$).

Since we chose the Supremum Theorem (or Axiom) as a focal point, let us mention that in the more recent elaboration of Rudin (1976), who takes back Dedekind's construction almost as it is, the Supremum Theorem is indeed the first consequence drawn from the theorem (explicitly attributed to Dedekind by Rudin) according to which any cut (A, B) of $\mathbb{R}$ is produced by an element of $\mathbb{R}$, this element being either the maximum of A or the minimum of B. Rudin's proof that any nonempty bounded subset E of $\mathbb{R}$ has a supremum consists in considering the cut (A, B), where $A = \{\alpha \in \mathsf{R} | \exists e \in E \text{ such that } \alpha < e\}$, the sought-after supremum being then the minimum of B.

A 'Set Theoretic' Viewpoint that Mobilizes Actual Infinity

Dedekind's project is to clear analytical proofs out of their recourse to geometrical arguments, but it does not prevent him from relying strongly on the (metaphorical) image of the number line to identify the object, in this case the cut, from which the construction will be built. As a matter of fact he makes no secret of it, and mentions from the outset this property of 'separability' of the line by its points, as the one he seeks to reproduce in the new set to be constructed.

One can also notice that sequences are absent, not only from his construction but also from the analytical theorems that he infers from it, and that are rather stated in terms of 'varying magnitudes' (i.e. functions). To our knowledge, the idea of sequence is evoked only one time, with the word 'successive' in the quotation "a variable magnitude x which passes through successive definite numerical values…" (Dedekind 1963a, p. 24).

Regarding the proofs of the Supremum Theorem or of the Theorem of increasing bounded functions, they require the mere introduction of a well-chosen cut, expressing in terms of cuts the inequalities in the hypotheses. The conclusion then follows almost directly from completeness 'à la Dedekind'. Arguments related to 'passages to the limit' or 'successive approximations' are absent, to such an extent that in the conceptualizations, prompted by the construction as much as by its exploitation, actual infinity is dominating, and potential infinity is solicited as little as possible. It is worthwhile recalling here that Dedekind (1888) is the first to have given a definition of an infinite set.

> 64. Definition. A system S is said to be infinite when it is similar[3] to a proper part of itself (32); in the contrary case, S is said to be a finite system. (Dedekind 1963b, p. 63).

Didactical Implications

Dedekind's construction of the set of real numbers highlights the role of order, and the insufficiency of density-in-itself, to warrant continuity. As shown in Durand-Guerrier (2016), this is a challenging issue in university mathematics education, especially because of the pervasiveness of the (figural) image of the line, which triggers moving directly from discreteness to continuity. Indeed, as Longo (1999, p. 403) writes: "The points are collected in the trace, which makes their individuality disappear. These points become evident again, as isolated points, when two lines cross each other." Durand-Guerrier (2016) proposes a teaching situation in which students are asked to prove the existence of a fixed-point for any increasing function from a finite segment of $\mathbb{N}$ into itself. Then, they are asked to

[3] According to Definition 32 from the same source, 'similar' means in a one-to-one correspondence (or *bijection*).

study the possible generalizations to increasing functions from the bounded interval [0; 1] into itself, first when [0; 1] is restricted to the set of finite decimal expansions, then when it is restricted to $\mathbb{Q}$, and finally when [0; 1] is in $\mathbb{R}$. The generalization is only possible in the setting of real numbers, thanks to $\mathbb{R}$ completeness. A classical proof relies on the construction of the cut of [0; 1] defined by $\{x \in [0;\ 1]; f(x) \geq x\}$. This subset is nonempty (it contains 0) and bounded above by 1; in $\mathbb{R}$, it has a least upper bound (lub) and this lub is a fixed point for the function. The teaching situation is aimed at fostering the understanding of the relationship between discreteness, density-in-itself and continuity, for an ordered set of numbers. In addition, we hypothesize that introducing the real number system through Dedekind's cuts and engaging students in proof and proving, with the main theorems presented above as objectives, would strengthen their capability related to existence proofs.

Real Numbers as Limit of Fundamental Sequences

During the same period when Dedekind built his system of real numbers, Cantor, working on trigonometric series, was developing his own theory of irrational numbers (Belna 1996, pp. 102–103). The first presentation of this theory was published in Cantor (1872), where he introduces the notion of *fundamental sequence*, in modern words "Cauchy sequence". According to Belna (1996, pp. 125–146), for Cantor, such sequences are to converge; and when they are not, he introduces a new number, that he calls an irrational number. It is defined by its position within $\mathbb{Q}$, the set of rational numbers, and by the possibility of extending the operations on $\mathbb{Q}$. The link with the numerical line is established a posteriori. Cantor poses that once an origin and a unit have been chosen, each point is defined by its abscissa. Cantor shows that in the case of an abscissa not being rational, there exists at least one fundamental sequence that determines this point, and he adds an axiom ascertaining that, for each magnitude, there exists a corresponding point on the line. Belna reminds us that Cantor did not introduce explicitly an equivalence relation between fundamental sequences, contrary to how this theory is presented in modern undergraduate and graduate textbooks, for reasons of both rigor and intelligibility. These reasons motivate our choice of accounting for this theory through the reworking proposed by Burrill (1967) and by Lelong-Ferrand and Arnaudiès (1977).

Construction of a Totally Ordered Field

Recall that a sequence $(a_n)_{n \in \mathbb{N}}$ is called 'fundamental', or a 'Cauchy sequence', if it satisfies the Cauchy criterion: $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $\forall n, m \in \mathbb{N}$, $n \geq m > N \Rightarrow |a_n - a_m| < \varepsilon$. Burrill denotes by C the set of all Cauchy sequences of rational numbers. He shows that these sequences are bounded, and that C is closed under term-by-term addition, subtraction, multiplication and division (provided the usual precautions for division).

He also shows that every convergent sequence of rational numbers is a Cauchy sequence, whereas the reciprocal is false. Now, the intuitive idea is that Cauchy sequences *should* converge. Hence the need for constructing a new set as an ordered field extension of $\mathbb{Q}$, in which it will be possible to prove that every Cauchy sequence is convergent. The idea is to regard two Cauchy sequences as equivalent *if and only if* their difference converges to zero. So an equivalence relation $\sim$ is defined on C by setting $(a_n) \sim (b_n) \Leftrightarrow (a_n - b_n) \overset{n\to\infty}{\rightarrow} 0$, and it is now a matter of identifying the set of real numbers with the quotient set $C/\sim$. It follows easily that if (a_n) converges to a in $\mathbb{Q}$, then $(b_n - a_n)$ converges to 0 *if and only if* (b_n) converges to a, so that in every equivalence class of C, either every sequence converges to the same rational limit, or none of them converges (to a rational limit). Hence a class of convergent sequences corresponds (canonically) to the rational being their shared limit, and this allows identifying $\mathbb{Q}$ with a subfield of $C/\sim$.

The next step is to define the order on $C/\sim$. To characterize a relation such as $(a_n) \leq (b_n)$ between classes, one should account for the class of $(b_n - a_n)$ being greater or equal to θ, the class of the identically zero sequence. Burrill defines a Cauchy sequence (a_n) to be 'non-negative' *if and only if* for every $\varepsilon > 0$, there exists a rank N such that for all $n > N$, $-\varepsilon < a_n$. He then shows that every Cauchy sequence (a_n) is either non-negative, or such that $(-a_n)$ is non-negative, and that the two will be simultaneously so *if and only if* they belong to θ. The proof rests decisively on the Cauchy criterion, and indeed the first result is not true in general for sequences that do not satisfy it. This is apparent in the formalism, since the statement "the sequence $(-a_n)$ is non-negative" is formalized by a universal statement, so that it cannot be the negation of the statement "(a_n) is non-negative". Because of that, the proof offers the opportunity to work on the definition of the Cauchy criterion and on its meaning. Burrill's proof uses *reductio ad absurdum* by establishing, in a concise way, a result that contradicts the Cauchy criterion satisfied by the given sequence.

To define order, Lelong-Ferrand and Arnaudiès (1977) define concurrently non-negative and non-positive sequences, and prove by a direct reasoning that the union of the two classes is C. We hypothesize that an analytical work on these two proofs would allow undergraduate students: first, to cope with the concept of Cauchy sequence in $\mathbb{Q}$, and in particular to develop reasoning about properties that are true 'save for a finite number of terms', i.e. true only from a certain rank on; second, to identify the role played by definitions in proof and proving; third, to work on the articulation between syntax and semantics, by acknowledging that certain forms can be non-equivalent syntactically—e.g. $(-a_n)$ is non-negative and the negation of (a_n) is non-negative—while being so semantically, because of the properties at issue and in connection with the alternation of quantifiers.

Now, the order on $C/\sim$ is defined by setting: $(a_n) \leq (b_n) \Leftrightarrow (b_n - a_n)$ is non-negative. Burrill shows: that this order is independent of the chosen representatives; that it is compatible with the operations, i.e. $\alpha \leq \beta \Rightarrow \alpha + \gamma \leq \beta + \gamma$; $\alpha \leq \beta$ and $0 \leq \gamma \Rightarrow \alpha \cdot \gamma \leq \beta \cdot \gamma$; and that it is a total order on $C/\sim$ (because either $(b_n - a_n)$ is non-negative, or $(-(b_n - a_n))$ is).

A Proof, via the Infimum Theorem, that the Field Thus Constructed Is Complete

It remains to show that the ordered field $(C/\sim, +, \times)$ is complete. Burrill writes: "Proof of completeness is the only nontrivial part of the construction" (Burrill 1967, p. 146). Surprisingly, he does not show, as one would expect, that every Cauchy sequence in $C/\sim$ is convergent. It is, as we will see, what is done by Lelong-Ferrand and Arnaudiès in line with Dedekind, whose proof of the 'continuity' of $\mathbb{R}$ amounts to proving that by repeating the construction, one does not get a new set. As far as Burrill is concerned, he gets out of the logic of his construction so to speak, in showing instead that the Infimum Theorem is verified in $C/\sim$. The arguments will rely on the ideas of decimal expansions and successive approximations.

For any $c \in \mathbb{Q}$, let $\varphi(c) \in C/\sim$ denotes the class of the constant sequence (c). A bounded above subset B of $C/\sim$ being given, Burrill constructs a sequence (c_t) of (finite) decimal expansions, such that for all $t \in \mathbb{N}$: $\varphi(c_t)$ is a lower bound for B and $\varphi\left(c_t + \frac{1}{10^t}\right)$ is not; $c_{t+1} = c_t + \frac{n_t}{10^{t+1}}$, where n_t is a integer between 0 and 9. Burrill then shows that the class of $C/\sim$ represented by (c_t) is the sought infimum for B. He concludes that the constructed set is a complete ordered field and that, because such a number system is unique as he established in his Chap. 7, it is the same ordered field that he had constructed in Chap. 6, using infinite decimal expansions.

Let us pause for a first reflection about this construction. The privileged objects are sequences, more precisely Cauchy sequences. We may certainly appraise the construction as being more abstract than Dedekind's one, the elements of the new set—equivalence classes of Cauchy sequences—being more complex, and certainly more difficult to relate to the intuitive support of the number line. Moreover, the understanding that $\mathbb{Q}$ is 'naturally included' in the constructed set is far less immediate than with cuts. The idea that we are 'completing' $\mathbb{Q}$ by adding missing elements is also more difficult to conceptualize, unless one is already well aware that in $\mathbb{Q}$, the non-convergent Cauchy sequences are these sequences whose terms are 'piling' at the edge of the 'holes' in $\mathbb{Q}$, and has grasped in what way the construction amounts to fill these holes in.

Burrill's proof of the Infimum Theorem obliges him to introduce auxiliary constructions through successive decimal approximations, in reference to the first construction of $\mathbb{R}$ he gave previously in his book (see below). But by a paradoxical fair return, these arguments allow Burrill: to reinject a form of numerical-geometric intuition in his approach; to display the approximations by which the passage to the limit will be driven; and hence to precede the recourse to actual infinity with a line of reasoning drawing on potential infinity, thus facilitating the involved conceptual leap. In our view, this illustrates the complementarity of the two approaches, to make sense out of the concept of real numbers. Dedekind's approach with cuts supports a deeper understanding of the concept of least upper bound and of its uses in Analysis, while Cantor's approach fosters the ideas of sequences, of successive approximations, and 'actualizes' the passage from potential infinity to actual infinity.

A Proof, via Cauchy Sequences, that the Set of Real Numbers Is Complete

Save some details, the construction of Lelong-Ferrand and Arnaudiès (1977) is the same as Burrill's one. The essential difference is in their proof of completeness, which consists in proving that Cauchy sequences in $C/\sim$ converge in $C/\sim$. It should be noted that from the outset, Lelong-Ferrand and Arnaudiès call 'numerical line' this set and designate it by $\mathbb{R}$. Having partitioned $\mathbb{R}$ in $\mathbb{R}+\cup\mathbb{R}$, via a definition of positivity (of Cauchy sequences) analogous to Burrill's one, the authors use the property $\alpha \in \mathbb{R}+ \Leftrightarrow -\alpha \in \mathbb{R}-$ to set $|\alpha| = \max\{\alpha, -\alpha\}$, thus extending from $\mathbb{Q}$ to $\mathbb{R}$ the absolute value and the associated distance, contrary to Dedekind who considers that 'continuity' must be dissociated from metric (Benis Sinaceur 2008a, pp. 38–39). The authors then feel the need to situate 'metrically' the rational numbers in $\mathbb{R}$ through approximations theorems. Among these, Lelong-Ferrand and Arnaudiès show that $\mathbb{R}$ verifies the Archimedes' axiom and that $\mathbb{Q}$ is dense in $\mathbb{R}$. Before proving completeness, they show that any Cauchy sequence of rational numbers converges to the real number it represents in the quotient set, as one would expect. From this argumentation, we will examine a 'passage to the limit' rather subtle, whose justification is not much detailed in their text.

Let $(y_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathbb{Q}$ representing $\beta \in \mathbb{R}$. The rational $\varepsilon \in 0$ being fixed, there exists $N(\varepsilon) \in \mathbb{N}$ such that $n > p > N(\varepsilon) \Rightarrow |y_n - y_p|\varepsilon$, i.e. $y_p - \varepsilon < y_n < y_p + \varepsilon$. Then the authors state that p being fixed, one sees "by the very definition of the order relation on $\mathbb{R}$" (op. cit., p. 17) that for all $p > N(\varepsilon)$, $y_p - \varepsilon \leq \beta \leq y_p + \varepsilon$, from what they infer that β is the limit of (y_n).

But since p, and therefore n, depend on ε, we may evaluate that the inequalities $y_p - \varepsilon \leq \beta \leq y_p + \varepsilon$ either are not necessarily true for all p, or are not true for all ε. Let us look more closely into this last step by coming back to the underlying definitions. For a fixed p depending on ε, saying that for each $n > p$, the inequalities $y_p - \varepsilon < y_n < y_p + \varepsilon$ hold in $\mathbb{Q}$ amounts to say that the sequences $(y_n - y_p + \varepsilon)_{n\in\mathbb{N}}$ and $(y_n - y_p - \varepsilon)_{n\in\mathbb{N}}$ are respectively non-negative and non-positive (keeping in mind that ε is here fixed and is not the ε intervening in the definition of positivity). If we denote by β the class of (y_n), y_p and ε being represented in $\mathbb{R}$ by themselves because they are rationals, we then have $0 \leq \beta - y_p + \varepsilon$ and $\beta - y_p - \varepsilon \leq 0$ by definition of the order in $\mathbb{R}$, i.e. that $y_p - \varepsilon \leq \beta \leq y_p + \varepsilon$. So we did show that for any $\varepsilon > 0$, there exists a $N(\varepsilon)$ such that for all $p > N(\varepsilon)$, $\left|y_p - \beta\right|\varepsilon$, this being the definition of convergence in $\mathbb{R}$ save for ε belonging to $\mathbb{Q}$ instead of $\mathbb{R}$. But it is not a problem because of the (proven) density of $\mathbb{Q}$ in $\mathbb{R}$.

The completeness of $\mathbb{R}$ is then established by Lelong-Ferrand and Arnaudiès. Let (x_n) be a Cauchy sequence of elements of $\mathbb{R}$. Their idea is to construct a sequence of rational numbers (y_n) sufficiently close to (x_n) by choosing, for each $n \in \mathbb{N}$, a rational y_n verifying $x_n - \frac{1}{n} < y_n < x_n + \frac{1}{n}$ (density of $\mathbb{Q}$ in $\mathbb{R}$). Using an argument similar to the one we just discussed, it is shown, in an otherwise relatively standard fashion, that (x_n) converges to β, the class of (y_n) in $\mathbb{R}$.

One may assess the complexity of the interplay between potential infinity and actual infinity in this context, where the conclusions about limits are obtained under the condition that the 'slippery' validity of the inequalities can ceaselessly be repelled further, 'at infinity'. And this is in addition to the need for coordinating two 'chases' at infinity: one for each object of the sequence (x_n), and the 'diagonal' one, so to speak, of the sequence itself. As far as the Supremum Theorem is concerned, Lelong-Ferrand and Arnaudiès prove it essentially as in Rudin, the authors having introduced cuts in the ensuing chapter and having shown that in $\mathbb{R}$, "For every cut (A, B) there exists a unique real number c verifying for all $a \in A$ and all $b \in B$, $a \leq c \leq b$."

Didactical Implications

In line with the contemporary structuralist methods of the twentieth century, the reworking of Cantor's theory by Burrill and by Lelong-Ferrand and Arnaudiès are modern 'formalizations' and by comparison, Cantor's development could be assessed as less rigorous. But this formal approach tends to hide the nature of the mathematical objects at play. By presenting the last two proofs of the completeness property, we have tried to show the necessity of going back to the very nature of the objects in play and to their properties, when constructing and producing proofs; and this underlines the crucial role played by the dialectic between syntax and semantics in proof and proving. This supports our claim that working on proofs with undergraduates may contribute to conceptualization. Providing empirical evidence for this conjecture is one of the goals of the didactical component to be developed within the current research project.

Real Numbers as Decimal Expansions

The two constructions that we have presented so far are rather theoretical, and weave very few connections with what students are experiencing at the secondary level, where decimal numbers are widely used in mathematics as well as in other subjects (e.g. physics or economy). So, for pragmatic reasons, it would seem rather natural to introduce the real numbers through their decimal expansions. An example of such an approach is given in Perrin (2005). In the introduction of the chapter devoted to real numbers, Perrin recalls that the set of rational numbers is an ordered field, and that its order is dense-in-itself, but states that this field is 'too small' and needs to be 'completed'. To support this, he shows that an integer that is not a perfect square has no square root in the set of the rational numbers. He then recalls that rational numbers have repeating decimal expansions, and introduces the idea of considering non-periodic decimal expansions as numbers, but claims the necessity of first building up the set of real numbers. This brings him to give an axiomatic definition of $\mathbb{R}$ as "The unique Archimedean ordered field that satisfies the property of adjacent sequences[4]: if two sequences (u_n) and (v_n) are adjacent, then they converge to a common limit" (Perrin 2005, pp. 97–98, our translation). The author gives an example to stress that this axiom does not hold in the subfield of rational numbers. He then shows that a decimal expansion defines a real number, thanks to the adjacent sequences axiom; conversely, he shows that any real number has a decimal expansion by introducing its nth decimal approximations, in excess and by default. In addition to this axiomatic definition of $\mathbb{R}$, Perrin proposes, in an appendix to the chapter in question, a construction of the set of real numbers relying on decimal expansions, which consists in doing the previous work upside down (op. cit., pp. 103–107). He defines a real number as being any *proper* decimal expansion (i.e. without a period of repeating nines) and introduces, using intervals, the definition of convergence for a sequence of real numbers. Given a real number x (a proper decimal expansion), he shows that the two sequences defined by the nth decimal approximations, respectively in excess and by default, are adjacent sequences in the set of rational numbers, and converge to x in the set of real numbers. He then shows that the property of adjacent sequences holds in the set of real numbers that he has just constructed.

Considering our focus on the links between the way objects are introduced and the related proofs, we will now present how Burrill (1967), in "Foundations of real numbers", chooses his definition of 'real number'. He bases it directly on the definition of integers, in such a way that "it will be unnecessary to engage in a preliminary discussion of the rational number system" (p. 75). His definition relies on "the usual way of visualizing real numbers […] in terms of their representation to the base 10" (p. 75). According to Burrill, real numbers can be viewed as chains $a_0\ a_1\ a_2\ a_3\ldots$ of integers, where the first integer a_0 is arbitrary and all others lie between 0 and 9. This motivates the way he develops the system of real numbers in a formal way (his wording), as we will see below, leaning solely on the properties of integers.

Let us now begin the formal development of the real number system. We continue the convention that I denotes both the set and the system of integers;

[4]Recall that two sequences (u_n) and (v_n) are said to be *adjacent* if (u_n) is increasing, (v_n) is decreasing, for each $n \in \mathbb{N}$, $u_n \leq v_n$ and finally, $(u_n - v_n)$ converges to zero.

as before I_ω signifies the set of all non-negative integers, and $I_9 = \{n \colon n \in I, 0 \leq n \leq 9\}$. Let $\Re$ be the collection of all functions f such that

(a) f is from I_w into I,
(b) $f(n) \in I_9$ if $n \geq 1$,
(c) for any $m \in I_\omega$, there is some $n \in I_\omega$ with $n > m$ and $f(n) \neq 9$.

A *real number* is defined to be a member of $\Re$. (Op. cit., p. 77).

He introduces the symbol θ for the function that maps all non-negative integers onto 0, and he denotes e^p the function that maps p onto 1, and all other non-negative integers onto 0, e^0 being 'the *unit*', simply written e. He defines the order as being the lexicographic order, giving rise to the relation < (less than); he proves that θ is different from e (Theorem 1), that the relation < is transitive in $\Re$ (Theorem 2), and that $\leq$ is a linear order for $\Re$ (Theorem 3).

He then shows that "*a non empty subset of* $\Re$ *that is bounded below has an infinum*", this of course amounting to completeness. In his proof (p. 79), he uses the following property of I: "in the set I of integers, every non-empty subset that is bounded below has a minimum", and proceeds inductively. We summarize the proof. Let S be a nonempty subset of $\Re$ bounded below by g (recall that g and elements of S are functions from I_ω into I). The set of images of 0 by the functions in S is a subset of I bounded below by $g(0)$, and hence has a minimum, which is denoted a_0. Burrill then defines recursively a sequence $a_0, a_1, a_2, \ldots, a_n, \ldots$, such that $a_0 \in I$ and for all $n \geq 1, a_n \in I_9$. He shows that the function h defined, for $n \in I_\omega$, by $h(n) = a_n$, is a real number, i.e. it satisfies the conditions (a), (b) and (c). For (a) and (b), it is direct because of the construction of h; for (c), Burrill proves it by contradiction, showing that it is not possible for h to "terminate by nines". He then proves by contradiction that h is a lower bound for S, and finally proves that no real number k greater than h is a lower bound for S.

Section "Real Numbers as Limit of Fundamental Sequences" of the chapter is devoted to the system that Burrill calls $\Re_\omega$, of those real numbers that "terminate in zeros" (pp. 80–84). It is defined as the infinite union over k of all the subsets $\Re_k$ of $\Re$, defined by $\Re_k = \{f \colon f \in \Re \text{ and for all } n > k, f(n) = 0\}$. He then introduces the kth approximating function f_k of a given real number f, f_k being in $\Re_k$, and he shows that $f = sup\{f_k \colon k \in I_\omega\}$.

Didactical Implications

The comparison between these two ways of constructing the set of real numbers through decimal expansions highlights the importance of the definitions being chosen for the objects, when moving to proofs. Both authors consider that the construction of $\mathbb{R}$ through decimal expansions is the most natural, but while Perrin remains within the usual aim of completing $\mathbb{Q}$, Burrill's primary construction of $\mathbb{R}$ is done directly from the integers, focusing on order and recursion and referring, for integers and their properties, to Peano's axioms. At the core of the proof of completeness by Perrin, we came across the *k*th approximations in excess and by default, that gives a central role to terminating decimals. This emphasizes the dyad terminating decimal/non-terminating decimal, which remains often implicit at the secondary schooling level. By using the axiom of adjacent sequences for completeness, this construction could be considered as a pragmatic match to the construction using Cauchy sequences. On the other hand, by focusing on order and the integers' properties, Burrill shows, in his system, that a real number is the supremum of the set of its *k*th approximations by default. To the extent that order is prevailing over metric considerations, his construction could be considered as a pragmatic match to Dedekind's construction. In our view, such an analysis highlights the different facets and uses of decimal expansions, and calls for being more precise and careful when tasks involving real numbers as decimal expansions are scrutinized, before being used in teaching whether intended for university students or for secondary level students.

Conclusion

We are interested in the opportunities offered by working on proofs to deepen the knowledge about mathematical objects, in relation with their (possible) definitions. Having brought back the main elements of Dedekind's and Cantor's definition of $\mathbb{R}$, we studied: first, the proofs of $\mathbb{R}$ completeness; second, the proofs of the *Supremum Theorem* from Dedekind and Rudin in the setting of *cuts*, and from Burrill and Lelong-Ferrand and Arnaudiès in the setting of *Cauchy sequences*. In addition, we proposed two contrasting constructions leaning on decimal expansions. By this, we want to stress the importance of the way objects are being defined towards molding the proofs, even when the approaches seem at first sight rather similar (e.g. defining real numbers as decimal expansions).

From this study, we retain the following main features:

- The proof of $\mathbb{R}$ completeness in Dedekind's construction is simple, and ensues directly from defining real numbers by cuts, with an explicit recourse to the density (with respect to order) of $\mathbb{Q}$. The proof of the Supremum Theorem then follows more or less directly. But as these proofs bring into play the sole *actual* infinity, they do not provide any pragmatic access to real numbers. The notion of cut appears as a powerful means for proving the existence of limits or of suprema without approaching their value, mainly because it rests on order, without resorting to distance. We may think that, in turn, this prevalence of the order relation over the metric does not foster the topological foundations, or generalizations, of Analysis, since the topology is then the one induced by order on the line, and its (potential) extensions (e.g. to the plane) are unclear.
- As far as Cauchy sequences are concerned, they provide a proof of $\mathbb{R}$ completeness that is complex, and requires juggling with several types of objects:

elements of $\mathbb{Q}$, Cauchy sequences of rational numbers, real numbers as equivalence classes, and so on. The lines of reasoning demand here a clever understanding of Cauchy sequences, with a tricky interplay between potential infinity and actual infinity. An important step leans on the fact that any real sequence can be approximated by a sequence of rational numbers. Specifying this idea to the real numbers themselves enables one to better situate, metrically and in due course topologically, real numbers with respect to rational numbers.

- The proof by Burrill, based on the construction of $\mathbb{R}$ starting from the integers, points out the possibility of shortcutting the construction of $\mathbb{Q}$, and of moving directly from integers to real numbers through order properties, at the core of the construction. By contrast, the construction through decimal expansions proposed by Perrin highlights in the construction itself the role played by decimal approximations and by the density in $\mathbb{R}$ of finite decimal expansions.
- From a didactic standpoint, we formulate the hypothesis that the construction through Dedekind's cuts and the related proofs foster a comprehension of completeness as a passage from density to continuity, when density is understood with respect to order, thus being intrinsic to $\mathbb{Q}$, in contrast with the (topological) density of $\mathbb{Q}$ in $\mathbb{R}$ which is external to $\mathbb{Q}$. In our view, this passage from density to continuity is the main obstacle to an adequate conceptualization of real numbers (Durand-Guerrier 2016), whereas the literature in didactics (e.g. Gravemeijer and Doorman 1999) has a tendency to focus on the passage from discreteness to continuity. On the other hand, Cauchy sequences foster the appropriation of the links between the idea of successive approximations, bringing potential infinity into play, and the notion of limit, leaning on actual infinity. We furthermore hypothesize that, to be operational, these two trends must be conjoined with an approach leaning on infinite decimal expansions: a real number is viewed as the supremum of the set of its decimal expansions rounded down to the nearest 10^{-n}, or as the limit of the (Cauchy) sequence formed by these expansions. Indeed, these expansions allow relating the idea of approximation, pertaining to distance, to the lexicographic order on expansions, which is their natural order, 'natural' as much geometrically as numerically or semiotically.

Acknowledgements This research project has been initiated as part of Denis Tanguay's residence at the *Institut Montpelliérain Alexander Grothendieck*, where he was a Visiting Professor upon the invitation of the *Centre national de la recherche scientifique* (CNRS, France).

References

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

Belna, J. P. (1996). *La notion de nombre chez Dedekind, Cantor, Frege*. Paris: Vrin.

Benis Sinaceur, H. (2008a). Nécessité et fécondité des définitions: Fondements de la théorie des nombres de Richard Dedekind (1931-1916). *Travaux de logique, 19,* 29–72.

Benis Sinaceur, H. (2008b). Note introductive à «Continuité et nombres irrationnels». In Dedekind, R., *La création des nombres*, 23–56. Paris: Vrin.

Bergé, A. (2010). Students' perceptions of the completeness property of the set of real numbers. *International Journal of Mathematical Education in Science and Technology, 41*(2), 217–227.

Burrill, C. W. (1967). *Foundations of real numbers*. New York: McGraw-Hill.

Cantor, G. (1872). Ueber die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen. *Mathematische Annalen, 5,* 123–132.

Dedekind, R. (1872). *Stetigkeit und irrationale Zahlen*. Braunschweig: Vieweg.

Dedekind, R. (1888). *Was sind und was sollen die Zahlen?* Braunschweig: Vieweg.

Dedekind, R. (1963a). Continuity and irrational numbers. In *Essays on the theory of numbers* (English translation of Dedekind (1872) by W.W. Beman). New York: Dover.

Dedekind, R. (1963b). The nature and meaning of numbers. In *Essays on the theory of Numbers* (English translation of Dedekind (1872) by W.W. Beman). New-York: Dover.

Durand-Guerrier, V. (2016). Conceptualization of the continuum, an educational challenge for undergraduate students. *International Journal for Research in Undergraduate Mathematics Education, 2*(3), 338–361.

Frege, G. (1971). *Écrits logiques et philosophiques* (Translation and Introduction from C. Imbert). Paris: Seuil.

Gravemeijer, K., & Doorman, M. (1999). Context problems in realistic mathematics education: A calculus course as an example. *Educational Studies in Mathematics, 39*(1), 111–129.

Lelong-Ferrand, J. (1964). *Les notions mathématiques de base dans l'enseignement du second degré*. Paris: Armand Colin.

Lelong-Ferrand, J., & Arnaudiès, J.-M. (1977). *Cours de Mathématiques*, *tome 2*, *Analyse*, 4e édition. Paris: Dunod Université.

Longo, G. (1999). The mathematical continuum: From intuition to logic. In J. Petitot et al. (Eds.), *Naturalizing phenomenology* (pp. 401–425). Standford: Stanford University Press.

Perrin, D. (2005). *Mathématiques d'école. Nombres, mesures et géométrie*. Paris: Cassini.

Rudin, W. (1976). *Principles of mathematical analysis* (3rd Ed.). New York: McGraw-Hill.

Vergnaud, G. (1990). La théorie des champs conceptuels. *Recherches en didactique des mathématiques, 10*(2), 133–170.

Chapter 3
Types of Epistemological Justifications, with Particular Reference to Complex Numbers

Guershon Harel

Abstract Epistemological justification is one of the essential constructs of DNR—a conceptual framework for the learning and teaching of mathematics. In this chapter, I distinguish among three types of epistemological justification:

(1) Sentential epistemological justification (SEJ). This refers to a situation when one is aware of how a definition, axiom, or proposition was born out of a need to resolve a problematic situation.
(2) Apodictic epistemological justification (AEJ). This pertains to the process of proving. It is when one views a particular logical implication, a → b, in causality, or explanatory, terms—how a causes b to happen. This can take place in two forms. One might observe a, asks what are its possible consequences, and finds out that b is a consequence of it. Or one might observe b, asks about its causes, and finds out that a is a cause of it.
(3) Meta epistemological justification (MEJ). This refers to a situation when one not only possesses SEJ and AEJ, but also he or she is aware of how the sentence or the implication came into being.

These three types will be illustrated with examples from the field of complex numbers.

Keywords Epistemological justification · Intellectual need · DNR-based instruction in mathematics

This chapter rests heavily on earlier publications dealing with DNR-based instruction in mathematics[1] (DNR, for short), a conceptual framework for the learning and teaching of mathematics (see, e.g., Harel 2008a, b, c, 2013a, b). Its

[1]*DNR* stands for the three foundational instructional principles of the framework, Duality, Necessity, and Repeated reasoning, discussed later in this paper.

G. Harel (✉)
Department of Mathematics, University of California, San Diego, USA
e-mail: harel@math.ucsd.edu

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_3

main goals are: (a) to offer a new classification for the construct of epistemological justification, and (b) to further analyze this construct. Its definition is inextricably linked to the construct of intellectual need (Harel 2013a) and rests on four of the eight DNR premises. They are:

Knowledge of Mathematics: Knowledge of mathematics consists of two related but different categories of knowledge: all the ways of understanding and ways of thinking that have been institutionalized throughout history.
Knowing: Knowing is a developmental process that proceeds through a continual tension between assimilation and accommodation, directed toward a (temporary) equilibrium.
Knowledge-Knowing Linkage: Any piece of knowledge humans know is an outcome of their resolution of a problematic situation.
Subjectivity: Any observations humans claim to have made are due to what their mental structure attributes to their environment.

The definitions of intellectual need epistemological justification are as follows:

> If K is a piece of knowledge possessed by an individual or community, then, by the *Knowing-Knowledge Linkage Premise* there exists a problematic situation S out of which K arose. S (as well as K) is subjective, by the *Subjectivity Premise*, in the sense that it is a perturbational state resulting from an individual's encounter with a situation that is incompatible with, or presents a problem that is unsolvable by, her or his current knowledge. Such a problematic situation S, prior to the construction of K, is referred to as an individual's *intellectual need:* S is the need to reach equilibrium by learning a new piece of knowledge. Thus, intellectual need has to do with disciplinary knowledge being created out of people's current knowledge through engagement in problematic situations conceived as such by them. One may experience S without succeeding to construct K. That is, intellectual need is only a necessary condition for constructing an intended piece of knowledge…. [If one constructs K and, in addition, is aware of] how K resolves S, we say that that person has constructed an *epistemological justification* for K... [Epistemological justification, thus, is a conscious relation between S and K] … It constitutes the geneses of mathematical knowledge—the perceived reasons for its birth in the eyes of the learner. (Harel 2013b, p. 122)

An individual's epistemological justification may not, and often does not, coincide with the historical epistemological justification. For example, a mathematician's epistemological justification for real analysis as a field is unlikely to initially be formed in the same way it was formed historically, which, according to Bressoud (1994), was intellectually necessitated from Fourier's solution to Laplace's equation, $\partial^2 z/\partial w^2 + \partial^2 z/\partial x^2 = 0$.

In Harel (2013b), I discussed in length the following five categories of intellectual need, along with their cognitive origins and their role in the learning and teaching of mathematics:

(1) *Need for certainty*. This is the need to prove, to remove doubts. One's certainty is achieved when one determines, by whatever means he or she deems appropriate, that an assertion is true.

(2) *Need for causality*. This is the need to explain—to determine a cause of a phenomenon, to understand what makes a phenomenon the way it is. This need does not refer to physical causality in some real-world situation being mathematically modeled, but to logical explanation within the mathematics itself.
(3) *Need for computation*. This need includes the need to quantify and to calculate values of quantities and relations among them by means of symbolic algebra.
(4) *Need for communication*. This consists of two reflexive needs: *the need for formulation*—the need to transform strings of spoken language into algebraic expressions—and the *need for formalization*—the need to externalize the exact meaning of ideas and concepts and the logical justification for arguments.
(5) *Need for structure*. This need includes the need to re-organize knowledge learned into a logical structure.

The goal of this chapter is to discuss categories of epistemological justification, in relation to this variety of intellectual need, in the context of complex numbers, one of the DNR-based curricula which has been investigated recently (Harel 2013b). I distinguish among three categories of epistemological justifications:

(1) *Sentential epistemological justification* (*SEJ*). This refers to a situation when one is aware of how a definition, axiom, or proposition was born out of a need to resolve a problematic situation. It is called so because it pertains to sentences with objective and logical meaning.
(2) *Apodictic epistemological justification* (*AEJ*). This pertains to the process of proving; hence, the term *apodictic*. It is when one views a particular logical implication, $\alpha \Rightarrow \beta$, in causality, or explanatory, terms—how α causes β to happen; that is how α explains the presence of β. This can take place in two forms. One might observe α, asks "What are its possible consequences?", and finds out that β is a consequence of it. Or one might observe β, asks, "What are its causes?", and finds out that α is a cause of it.
(3) *Meta epistemological justification* (*MEJ*). This refers to a situation when one not only understands that a sentence is a resolution to a problematic situation or views an implication in explanatory terms, but also he or she is aware of how the sentence or the implication came into being.

The three categories emerged from analyses revolving around the questions whether mathematics instruction attends to students' intellectual need when introducing mathematical statements and proofs, and whether students are aware of such a need when they experience it. In the rest of this chapter, I will illustrate these types of epistemological justification with examples in the context of complex numbers. These are mere illustrations, not claims of existence. Their potential value is that they might be used as indicators and conceptual labels of empirical observations, established through accepted research methodologies, in the sense of Corbin and Strauss (1990), and possibly as initial hypothetical models for student reasoning, again to be substantiated empirically.

For the sake of completeness, I outline the general steps in the flow of lesson that is structured around the construct of intellectual need:

(1) Recognize what constitutes an intellectual need for a particular population of students, relative to a particular subject (in our case complex numbers).
(2) Translate this need into of a set of problematic situations which the students can potentially understand.
(3) Help students elicit the concepts from their solutions to these problems.

Figure 3.1 depicts the DNR elements discussed in this chapter. For a fuller discussion of these and other DNR elements, see Harel (2008a, b, c, 2013a, b, in press).

Before proceeding, a word on the relationship between epistemological justification and reasoning and proof (the subject of this volume) is in order. Brousseau (1997), in his characterization of the work of the mathematician versus that of the teachers, writes:

> Before communicating what she thinks she has discovered, a mathematician must first identify it. It is not easy, within the maze of thoughts, to distinguish what has potential of becoming new knowledge of interest to others …. In addition, all irrelevant reflections must be suppressed …. *One must conceal the reasons which led her in these directions and the personal influences which guided success. …*
>
> The teacher's work is to some extent the opposite of the [mathematician's]; *she must produce a recontextualization and a repersonalization of the knowledge …* (emphases added; pp. 21–22)

The notion of epistemological justification and its various instantiations, as described in this chapter, pour content into the processes of "recontextualization and repersonalization of knowledge" referred to by Brousseau, in that they can potentially provide the teacher with cues as to the constituents of reasoning and proof she advised to attend to in her instruction.

In addition, there is an inherent pedagogical inconsistency in instruction that emphasizes rigor without attention to the origin and need of that rigor. For on the one hand this type of instruction demands justification of conjectures and assertions, but ignores the need to justify their origins. As a consequence, students feel aliens in knowledge construction. This and several other papers aim at raising an awareness of this inconsistency, in hope that other scholars would conduct studies of the efficacy of the ideas surrounding the notion of epistemological justification.

Instances of Sentential Epistemological Justification (SEJ)

The first part of this section discusses fleetingly the historical development of complex numbers. (For a fuller account of this history, see Tignol 1980.) This development begins in a form of a need for computation—to find a solution to the cubic equation. The 16th Century mathematicians partially resolved this need by first finding a formula to the equation, $x^3 + mx = n$. The solution was discovered first by del Ferro and again by Tartaglia. In both cases, no justification was provided

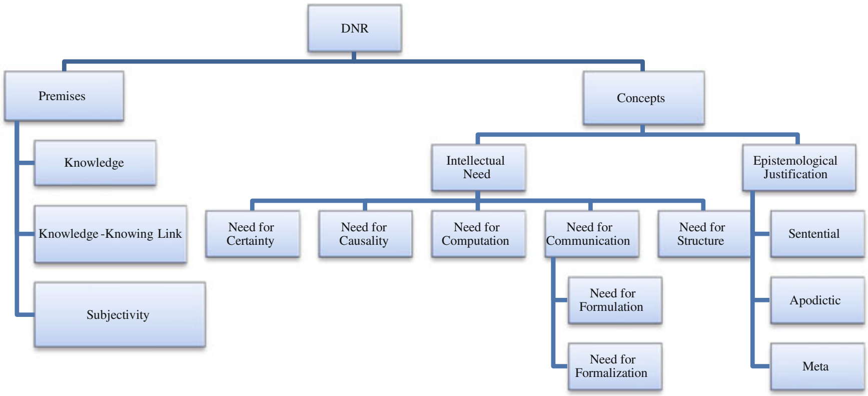


Fig. 3.1 DNR elements pertaining to the concern of this chapter

to the solution. A combination of the need for computation—to find a formula for the general cubic equation, $x^3 + ax^2 + bx + c = 0$—and the needs for certainty and causality—to prove the formula, possibly in an explanatory manner—led Cardano to the resolution of both needs. In modern terms, his formula can be stated as follows:

$$y = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2}}$$

where

$$y = x + \frac{a}{3}, \quad p = b - \frac{a^2}{3}, \quad \text{and} \quad q = c - \frac{a}{3}b + 2\left(\frac{a}{3}\right)^3$$

As the mathematicians of the time looked into this new result, they encountered baffling behaviors which raised doubts in the validity of the formula, triggering the need for certainty.

First, the cubic formula, unlike the quadratic formula, which was known at the time, did not yield all the roots.

Second, the formula often yields complicated expressions for simple roots. For example, while $x = 1$ is a solution—the only solution—to the equation, $x^3 + x = 2$, the cubic formula yields the complex expression, $x = \sqrt[3]{1 + \frac{2}{3}\sqrt{\frac{7}{3}}} + \sqrt[3]{1 - \frac{2}{3}\sqrt{\frac{7}{3}}}$ as a solution.

Third, and most perplexing, behavior of the cubic formula is that in certain cases the formula yields meaningless expressions when "real" roots are known. For example, for $x^3 = 15x + 4$, the formula yields, $x = \sqrt[3]{2 + \sqrt{-121}} + \sqrt[3]{2 - \sqrt{-121}}$, where it can easily be seen that $x = 4$ is a root.

Fourth, and of particular importance, when simplification procedures were applied to the latter "meaningless" expressions involving $a + b\sqrt{-1}$, treating them as if they were meaningful, the cubic-root addends often yielded expected results. For example, the addends in $x = \sqrt[3]{2 + \sqrt{-121}} + \sqrt[3]{2 - \sqrt{-121}}$, turned out to be, respectively, $2 + \sqrt{-1}$ and $2 - \sqrt{-1}$, which, in turn yields the solution, $x = \left(2 + \sqrt{-1}\right) + \left(2 - \sqrt{-1}\right) = 4$, as expected. While this work (by Bombelli) provided some assurance about the validity of the cubic formula, it engendered further puzzlement: How is it possible that meaningless expressions turn under legitimate manipulations into meaningful results?

These behaviors led to further investigations into the meanings and roles of the expressions, $a + b\sqrt{-1}$. When these meanings and roles were found, these expressions received the status of numbers and were defined as such.

This account is an example of SEJ. It manifests an understanding of the definition (a sentence) of complex numbers as a consequence of certain intellectual needs, particularly the need for computation. Clearly, one may understand complex number possessing its SEJ.

The development of complex numbers provides another example of SEJ, this time for the emergence of crucially important theorem—the Fundamental Theorem of Algebra (FTA). The emergence of complex numbers as an extension of the field of real numbers raised the question, Is there a need for further extensions? That is to say, does the complex field contain all the solutions to any polynomial equation? This question, which saliently belongs to both the need for computation and the need for structure, was answered affirmatively by Gauss. Thus, understanding the FTA as a resolution to these needs is another example of SEJ. Unfortunately, seldom do abstract algebra students possess such an understanding, as our experience shows. On multiple occasions, I asked students who completed successfully an abstract algebra course, what is fundamental about the FTA? And what question does the FTA answers? In all cases, none of the students was able to answer these questions.

Instances of Apodictic Epistemological Justification (AEJ)

In this section, I illustrate apodictic epistemological justification by restructuring Cardano's proof for the cubic formula as to delineate its underlying ideas, rendering it a causal proof. In fact, in Cardano's proof the solution formula is only a sufficient condition for the cubic equation. The following proof, formulated in modern symbolism, is a slight modification of Cardano's proof, in that it uses the identity, $(u + v)^3 = u^3 + 3u^2v + 3v^2u + v^3$, thereby making Cardano's solution formula equivalent to the cubic equation (Harel 2013a):

To solve the equation

$$x^3 + px + q = 0 \tag{3.1}$$

where $p, q \neq 0$, let $x = u + v$.

By cubing both sides of Eq. (3.1), expanding, and factoring uv, we get

$$x^3 - 3uvx - \left(u^3 + v^3\right) = 0 \tag{3.2}$$

Equations (3.1) and (3.2) are equivalent if and only if

$$\begin{cases} uv = -\frac{p}{3} \\ u^3 + v^3 = -q \end{cases} \tag{3.3}$$

System (3.3) is equivalent to the system,

$$\begin{cases} v = -\frac{p}{3u} \\ 3^3(u^3)^2 + 3^3 qu^3 - p^3 = 0 \end{cases} \tag{3.4}$$

The quadratic equation in system (3.4) is equivalent to

$$u = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2}} \quad \text{or} \quad u = \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2}}. \tag{3.5}$$

By symmetry between the variables u and v in system (3.3) and the second equation of the same system, we get that system (3.4) is equivalent to

$$u = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2}} \quad \text{and} \quad v = \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2}}$$

Hence, x is a solution to Eq. (3.1) if and only if

$$x = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2}}.$$

This proof rests on the ingenious idea of conceiving of a solution x as a sum of two numbers, $u + v$, for the purpose of reducing the cubic equation into a system of equations, which, in turn, is reducible to a quadratic equation. The critical questions relevant to AEJ is: How to intellectually necessitate the various elements of this proof to the students for whom the unit is intended, making its various steps explanatory in the eyes of the student?

In what follows I describe briefly the process I implemented in a teaching experiment with secondary mathematics teachers, which aimed at answering this question. The process consists of a sequence of lessons encapsulated into a sequence of the following four perturbation-resolution pairs:

Perturbation 1: How to solve systems of equations of a new kind?

The first lesson begins with word problems whose solution requires the solution of systems of the form

$$\begin{cases} uv = P \\ u^3 + v^3 = Q \end{cases} \qquad (*)$$

Typically the teacher participants are familiar with simple systems, mostly 2×2 and 3×3 linear systems, and so the only difference here is the form of the system. As can easily be seen, system (*) is of the same form as that of system (3.3) in the proof of the cubic formula presented above.

Resolution: Reduce the system into a quadratic equation (RQE).

The participants, after some struggle reduced this system into the quadratic equation, $(u^3)^2 - Qu^3 + P^3 = 0$.

Perturbation 2: What to do if the system is not reducible to a quadratic equation?

As the participants practiced the RQE technique on a family of systems involving products and cubes of unknowns, they encountered one system for which the technique leads to an irreducible 6-degree polynomial equation. The system was of form, with $P = -2$ and $Q = 32$.

$$\begin{cases} \frac{uv}{u+v} = P \\ u^3 + v^3 = Q \end{cases} \qquad (**)$$

Resolution: Reduce the system into a cubic equation (RCE).

After failed attempts to solve this system, the participants were reminded (in Lesson 2) that the three expressions, uv, $u + v$, and $u^3 + v^3$, in system (**) were parts of the identity, $(u+v)^3 = u^3 + 3u^2v + 3uv^2 + v^3$. In turn, this led to a reduction of system (**) into the cubic equation, $(u+v)^3 - 3P(u+v)^2 - Q = 0$. The technique developed was then applied to a series of systems, all were designed to be reducible to "easily solvable" cubic equations, i.e., equations that can be solved by finding one root of the equation through trial and error or by the Rational Root Theorem, and then finding the rest of the roots by the Division Theorem.

Perturbation 3: What to do if the RCE leads to a cubic equation that is not "easily solvable?"

The participants were then introduced to the following system:

$$\begin{cases} uv(u+v) = 8 \\ u^3 + v^3 - 2u - 2v = -31 \end{cases} \qquad (***)$$

They successfully used the RCE technique to reduce this system into the cubic equation, $x^3 - 2x + 7 = 0$, but since the equation has no rational roots, they were not able to solve it by trial and error or the Rational Root Theorem. This difficulty, in turn, led to the question: How to solve cubic equations? That is: Is there a solution formula for cubic equations, as in the case of quadratic equations?

Resolution: Develop a formula for cubic equations; first focus on those without the second term.

Since the problem at hand is solving the equation, $x^3 - 2x + 7 = 0$, Lesson 3 retains its focus on equations of the same form; namely: $x^3 + Ax + B = 0$. By revisiting the solutions to systems (*) and (**) along with extensive discussions and reflections on the techniques used this far, the participants successfully develop a solution to the equation. The solution amounts to a generic proof of Cardano's formula.

Perturbation 4: What to do with cubic equations with a second term?

In Lesson 4, the participants encountered an obstacle: neither the cubic formula known this far nor the re-application of the technique applied to develop it (i.e., the combination of RQE and RCE techniques) are successful for solving equations involving a second term.

Resolution: Reduce equations with a second term into ones without a second term.

Lesson 5 begins by revisiting the quadratic equation, $x^2 + Bx + C = 0$, and showing the participants how the change of variable $x = y + \frac{-B}{2}$ reduces the equation to one without the second term, and how, similarly, the change of variable $x = y + \frac{-B}{3}$ in the cubic equation $x^3 + Bx^2 + Cx + D = 0$ leads to a cubic equation without the second term, resulting in a cubic equation of a desired form: $x^3 + Ax + B = 0$. With this knowledge in hand, the cubic formula for the most general cubic equation was derived by the participants.

Figure 3.2 represents this sequence perturbation-resolution pairs, rendering Cardano's proof into a corresponding sequence of apodictic epistemological justifications.[2]

[2]This sequence of perturbation-resolution pairs was finalized after several iterations of teaching experiments. In Harel (2013b), I describe the role the instructor played in necessitating for the students certain paths to resolve perturbations.

***Perturbation 1*: How to solve systems of equations of a new kind?**

$$\begin{cases} uv = P \\ u^3 + v^3 = Q \end{cases}$$

Resolution: Reduce the system into a quadratic equation.

$$\left(u^3\right)^2 - Qu^3 + P^3 = 0$$

***Perturbation 2*: What if the system is not reducible to a quadratic equation?**

$$\begin{cases} \dfrac{uv}{u+v} = P \\ u^3 + v^3 = Q \end{cases}$$

Resolution: Reduce the system into a cubic equation.

$$(u+v)^3 - 3P(u+v)^2 - Q = 0$$

***Perturbation 3*: What if the reduction leads to a cubic equation that is not "easily solvable"?**

$$\begin{cases} uv(u+v) = 8 \\ u^3 + v^3 - 2u - 2v = -31 \end{cases}$$

Resolution: Use these reduction ideas to develop a formula for cubic equations w/out the 2nd term.

$$x^3 + Bx + C = 0$$

***Perturbation 4*: What to do with cubic equations with a second term?**

$$x^3 + Bx^2 + Cx + D = 0$$

Resolution: Reduce into equations without a second term, using change of variable.

Fig. 3.2 Cardano's proof as a sequence of apodictic epistemological justifications

Instances of Meta Epistemological Justification (MEJ)

MEJ is a subtle cognitive phenomenon, definitely more so than SEJ and AEJ. In all of the examples I discussed in the previous section, one might develop an understanding of how the definition of complex number was born out of the five intellectual needs that emerged during the development of the cubic formula, leading up to the definition of "complex number".[3] But, in this process one might not know the actual struggle—the pitfalls and insights—the community went through to arrive into satisfactory, institutionalized resolutions of these needs. Similarly, one might understand the re-presentation of Cardano's proof as a causal proof, but he or she may not be aware of the essential mental processes the author of this modified proof carried out in the process of obtaining it. Through exerted effort of reflection, the

[3]For example: The need for computation for the development of the cubic formula; the needs for causality, communication, and structure for the resolution of the puzzling aspects of the formula; and the need for certainty for its proof.

individual might gain an awareness of such a process, whereby gaining MEJ for the respective knowledge. For example, in the original Cardano's proof the solution formula is only a sufficient condition for the cubic equation. By recognizing the existence of this gap and comparing the original proof to the modified proof, one might observe the following important difference: The identity $(u+v)^3 = u^3 + 3u^2v + 3v^2u + v^3$ was generated in the new proof precisely to make the solution formula a *true formula*—that is, to make the cubic equation equivalent to the cubic formula, in that a number satisfies one if and only if it satisfies the other.

Contrast with Current Instructional Treatment of Complex Numbers

Current instructional treatment of mathematics in general and of complex numbers in particular for high-school students is devoid of epistemological justification. Students are simply given the definition of complex number as an expression of the form: $a + b\sqrt{-1}$, where a and b are real numbers. They then are taught how to apply arithmetic operations to such expressions. The following conversation between the author of this chapter and undergraduate students who learned complex number in their high schools demonstrates the impoverished understanding of complex numbers students have. These students were participants on a teaching experiment involving complex numbers (see Harel 2013b). When the $a + b\sqrt{-1}$ first appeared in the development described earlier, the question about the meaning of such expressions was raised. Then the conversation along ensued:

Students	$\sqrt{-1}$ is the complex number i.
Teacher	What does this mean?
Students	It means that $i^2 = -1$
Teacher	So we define i to be a number such that $i^2 = -1$. This is fine, but for what purpose?
Students	To solve the equation, $x^2 + 1 = 0$.
Teacher	That is true. But consider this: We create a new number, i, to turn an equation with no solution into one with a solution. Why then don't we do the same for other equations, such as $x + 1 = x + 2$ or its equivalents? Why don't we create numbers for such equations to turn them into equations with a solution?
Students	That was what we were told in school. Really, why don't we?—Why do we treat $x^2 + 1 = 0$ differently from all other equations that don't have solutions?

This was productive dialogue for it fostered a need with the students to better understand how complex numbers came about and the role they serve in mathematics. The dialogue also demonstrates that the way complex numbers are traditionally introduced in elementary algebra is abrupt and rather contrived.

Concluding Remark

I have defined in this chapter three categories of epistemological justification, *sentential epistemological justification (SEJ), apodictic epistemological justification (AEJ)* and *meta epistemological justification (MEJ).* While SEJ and MEJ are new contributions, AEJ is akin to Leron's (1983, 1985) idea of *structural proof,* Steiner's (1978) debate on *mathematical explanation*, and Hanna's (1990, following Steiner) distinction between *proofs that prove and proofs that explain*.

The pedagogical goal of structural proof, as stated by Leron, is to make the learner *aware* of the ideas hidden in the traditional, linear presentation of proofs. The source "structure" here intends to convey the essential act of the method: the process of *restructuring* a linear presentation of a proof, as is commonly presented in textbooks or research papers, to a multi-dimensional presentation that conveys possible thought processes involved in the construction of a proof, how one comes up with a piece of knowledge K to resolve a problematic situation S. Leron, and others (e.g., Alibert and Thomas 1991), discusses many excellent examples to illustrate the application of the method in converting linear proofs into a structural proof.

The notion of causal, or explanatory, proof has been discussed in the literature of the philosophy of mathematic (Steiner 1978) and of mathematics education (Hanna 1990; Hanna et al. 2010), and both can be traced to Aristotle's definition of scientific knowledge. The contribution of our notion of AEJ relative to these works is in the specific and elaborate connection of epistemological justification to a variety of intellectual needs, in its emphasis on the subjective learner, and its emphasis on proof production, not only proof comprehension.

References

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

Bressoud, D. (1994). *Radical approach to real analysis*. Washington, D.C.: Mathematical Association of America.

Brousseau, G. (1997). *Theory of didactical situations in mathematics*. Dordrecht: Kluwer.

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology, 13*(1), 3–21.

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

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

Harel, G. (2008a). DNR Perspective on mathematics curriculum and instruction: Focus on proving, part I. *ZDM—The International Journal on Mathematics Education*, *40,* 487–500.

Harel, G. (2008b). DNR Perspective on mathematics curriculum and instruction, part II. *ZDM—The International Journal on Mathematics Education, 40,* 893–907.

Harel, G. (2008c). What is mathematics? A pedagogical answer to a philosophical question. In B. Gold & R. Simons (Eds.), *Proof and other dilemmas: Mathematics and philosophy* (pp. 265–290). Washington, D.C.: Mathematical Association of America.

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

Harel, G. (2013b). DNR-based curricula: The case of complex numbers. *Journal of Humanistic Mathematics, 3*(2), 2–61.

Leron, U. (1983). Structuring mathematical proofs. *The American Mathematical Monthly, 90*(3), 174–185.

Leron, U. (1985). Heuristic presentations: The role of structuring. *For the Learning of Mathematics, 5,* 7–13.

Steiner, M. (1978). Mathematical explanation. *Philosophical Studies, 34*(2), 135–151.

Tignol, J. (1980). *Galois' theory of algebraic equation*. Essex, England: Longman Scientific & Technical.

Chapter 4
Mathematical Argumentation in Elementary Teacher Education: The Key Role of the Cultural Analysis of the Content

Paolo Boero, Giuseppina Fenaroli and Elda Guala

Abstract The problem dealt with in this chapter concerns how to prepare prospective elementary teachers to develop students' argumentative skills in school, in spite of difficulties deriving from present school culture and past teacher education in Italy. The salient features of a course on mathematical argumentation, aimed at making prospective elementary teachers free from those influences and enable them to perform autonomous professional choices, are described. The development of the competence of Cultural Analysis of the Content (CAC) is motivated as a condition for teachers' professional autonomy. Specific educational choices and some results concerning the development of participants' CAC in the course at stake are presented and discussed.

Keywords Mathematical argumentation · Elementary teacher education
Cultural analysis of the content · Argumentation and proof

Introduction

The necessity of promoting mathematical reasoning and proof at every school level is widely acknowledged now, and in many countries national programs and guidelines for curricula take it into account. In the Italian Guidelines for Curriculum (MIUR 2012), two "goals for the development of competencies" out of eleven for grade V (10-year-olds) concern mathematical argumentation. One of the main problems to be tackled in order to attain such goals concerns teacher education, given the school cultural environment in which teachers will have to teach. In Italy, great difficulties in promoting these aims at the elementary school level (grades I–V)

P. Boero (✉) · G. Fenaroli
DIMA and DISFOR, Genoa University, Genoa, Italy
e-mail: boero@dima.unige.it

E. Guala
DIMA, Genoa University, Genoa, Italy

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_4

derive from two factors: misunderstandings among teachers about the nature of mathematical reasoning; and the prejudice existing in present school culture about children's lack of capacity to develop genuine mathematical reasoning (particularly as concerns mathematical argumentation and proof).

These are inter-related factors, rooted in the New Mathematics movement, on the one side, and in the interpretation of some ideas from Piaget, on the other. In Italy (differently from other countries, like France), New Mathematics was never implemented in national programs; however New Mathematics was (during the sixties and the early seventies) an object of a massive national teacher training program. It also inspired some parts of widely adopted textbooks. The failure of related teaching experiences was usually interpreted by teachers as a consequence of the abstract nature of mathematical reasoning, considered to be out of the reach of elementary school students (Pellerey 1989). Note that, till the end of the last century, elementary school teacher education took place in a 4-year specialized high school, where teaching of Piaget's theory (the only theory taught from the field of developmental psychology) concerned a few rough ideas about the stages of intellectual development. As a result, in many Italian grade I–V classes, mathematics is taught as a set of notions concerning numbers and geometrical entities, and as practical recipes about how to solve a standardized set of word problems, contrary to the National Guidelines for the Curriculum.

Since the end of the last century, elementary teacher preparation has been reformed. Prospective teachers have to complete a 4-year (after 2010, 5-year) master's course in Sciences of Elementary Education, including 22 credits in Mathematics and Mathematics Education. In spite of these changes, however, elementary teacher education still faces several difficulties in the case of mathematical argumentation: besides the school culture and traditions (see above), most of our prospective teachers were not good in mathematics in secondary school, and frequently were even afraid of it. Similar difficulties exist in other countries. Stylianides, Stylianides and Shilling—Traina (2013) consider obstacles inherent in prospective elementary teachers' "weak mathematical (subject matter) knowledge about reasoning-and-proving, and counterproductive beliefs about its teaching" (p. 1563). One of the problems that we have tried to solve in our research concerns how to enable prospective teachers to identify and master the skills related to mathematical argumentation (see later), which may be accessible to elementary school students and whose educational relevance goes beyond the boundaries of mathematics. For this reason, we tried to develop the competence of Cultural (epistemological, historical and anthropological) Analysis of the Content (CAC) (Boero and Guala 2008) in teacher education courses, in particular in the case of mathematical argumentation.

Our chapter addresses the framing, implementation, and a partial account of the results of this implementation, of a course on mathematical argumentation. The course belongs to the set of mathematics and mathematics education integrated courses for elementary teacher education at our university; they are planned and taught by a team of mathematics educators including the authors of this chapter. Two inter-related main aims of the course (corresponding to the negative factors described above) were: making prospective elementary school teachers aware of different, historically legitimate ways of proving in mathematics, and their salient features; and dispelling their doubts (and the prejudice resulting from present school culture) about the possibility that elementary school students may have access to mathematical argumentation (proving in particular). In this chapter we will provide some evidence about the effectiveness of the performed theoretical and methodological choices we made to achieve those aims through the development of participants' CAC competence. We may observe that, in spite of their specific motivation in the Italian context, the aims of the course can be of broader interest, especially in those countries where guidelines for curricula stress the importance of developing mathematical argumentation as early as the elementary school.

Theoretical Framework

Research for Innovation

The development of our courses for elementary teacher education at the university level was strongly influenced by our research experience in the design and implementation of long term teaching experiments in elementary and lower secondary school. They conformed to the paradigm of research for innovation, shared by several researchers in Italy (see Arzarello and Bartolini Bussi 1998). In our case, we planned our first teacher education courses according to Schoenfeld's ideas (Schoenfeld 1994) on the crucial role of problem solving at an adult level in teacher preparation. Then we moved to a gradual enrichment of theoretical perspectives resulting from the necessity of tackling problems emerging from the analysis of the implemented courses, and from the needs of professional development, taking into account the situation of elementary school teaching of mathematics in Italy as described earlier. The CAC construct (Boero and Guala 2008) became a key reference for the design of our courses. In the specific case of the course addressed, after 2013, to develop students' competencies related to mathematical argumentation (according to the National Guidelines for Curriculum), we realized that we needed further theoretical tools to plan our courses: in particular, we adapted Habermas' elaboration on rationality in order to compare different ways of proving (Boero 2006) and argumentation in different cultural domains (Guala and Boero 2017). The development of the CAC competence helped to focus prospective teachers' attention on the anthropological and historical dimension of argumentation.

Cultural Analysis of the Content (CAC) Competence

We assume that teachers need to develop a competence of CAC, important for the teacher's cultural autonomy (see Boero and Guala 2008; Boero et al. 2014). CAC competence is different from theorized kinds of teacher knowledge like Pedagogical Content Knowledge (Shulman 1986) or Mathematical Knowledge for Teaching (Ball et al. 2008): first, CAC is more focused on the capacity of performing an activity relevant for professional choices, rather than on specific knowledge needed for them; second, it concerns cultural aspects, which were only partially considered in research on teachers' knowledge and conceptions, particularly in the case of proof and proving. Regarding teachers' knowledge about proof, Stylianides and Ball (2008) considered some epistemological characteristics of proof (logical-linguistic aspects) and of related tasks suitable to engage students in proving, while Knuth (2002) dealt with the functions of proof and related teachers' conceptions. Anthropological and historical-epistemological aspects of argumentation and proof (the relationships between argumentation in different domains and mathematical proof, the historical evolution of criteria of legitimacy for mathematical proof, etc.) are scarcely considered in the literature on teachers' knowledge.

We think that the competence of dealing also with those aspects of argumentation (and proof in particular) is crucial for the teacher for different reasons. First of all (and this is particularly true in Italy, for the reasons indicated in the Introduction) the teacher must be aware of the different ways of dealing with proof through the history. When planning didactic sequences and situations, and when analyzing students' productions, the historical analysis of proof and proving suggests a wide set of possible approaches to mathematical argumentation in the classroom and possible historical references for difficulties met by students (see Radford et al. 2000 for general considerations, Grabiner 2012, and Guala and Boero 2017 for specific issues concerning proof and proving). Second, the relevance of mathematical argumentation stems not only from its role in mathematical enculturation, but also from its relationships with general enculturation and extra-school culture (see Siu 2012, for an intercultural perspective). Mathematical argumentation is a key component of mathematical activities and, at the same time, strictly related to argumentative competences in other domains and in everyday life. The development of the CAC competence for mathematical argumentation (and proof in particular) was the main goal of the course; it was gradually made explicit for participants during the course.

Habermas' Construct of Rationality

As adapted to mathematics education in previous research (Boero 2006), Habermas' construct of rationality offers an analytical tool to compare proving in different mathematical domains (Guala and Boero 2017) and, given a mathematical domain, in different historical periods and cultures (cf. Durand-Guerrier et al. 2012). It also points out some general, *common* requirements for *different* kinds of mathematical reasoning and *proving* (as *rational behavior*).

According to Habermas (1998), rationality in discursive practices consists of three inter-related components:

- epistemic rationality, as *rationality of knowing* (consciously moving from knowing something, to knowing why it is true in a given cultural context);
- teleological rationality, as *rationality of doing* (consciously moving from producing strategies to achieve the aim of the activity, to evaluating those strategies in the perspective of reinvesting some of them in the future);
- and communicative rationality, as *rationality of communicating* (through the conscious search for adequate means to reach the interlocutor's understanding).

The Habermas' construct of rationality was used, in this course, as a tool for planning and managing our teaching at the university level; it did not become an explicit tool for prospective teachers (differently from a parallel course for secondary school teacher education—see Guala and Boero 2017). The reasons for this choice were the limited available time, together with the lack of a sufficient prospective teachers' background in history and epistemology of mathematics. We will discuss this choice in the last section of this chapter.

Argumentation and Toulmin's Model

While adapting Habermas' construct of rationality to mathematical activity (in particular, to proving), Toulmin's model of argumentation was identified as a valuable analytical tool for epistemic rationality (see Boero et al. 2010). According to Tsamir's (2008) criteria for introducing theoretical tools in teacher education, Toulmin's model of argumentation (Toulmin 2008) was explicitly taught in our course as a tool for identifying argumentation and dealing with it in task design and in the analysis of participants' and students' argumentative productions. In particular, in the course at stake we defined "argumentation" as an oral or written text organized around one nuclear argumentative step, or some enchained nuclear steps. Inspired by Toulmin's model, a nuclear argumentative step may be represented this way:

Data -------------------- Claim

|

Warrant

Among the advantages of this model (and its schematic representation) we would like to point out how it allows to identify lacking parts in students' argumentation, to classify different kinds of mathematical argumentation according to the nature of warrants (e.g., empirical vs theoretical warrants) and to design argumentative tasks (e.g., tasks asking for the identification of an inappropriate warrant). Some examples will be provided later.

Other Theoretical Elaborations

Other theoretical elaborations of different origins (made locally coherent for our purposes) served as references for specific aspects of our work. Some of them had been objects of explicit teaching in the other courses taught by us. They were: Vygotsky's elaboration about the role of language, the teacher's role in the students' *zone of proximal development*, and the *everyday concepts—scientific concepts* dialectics; and Vergnaud's *theory of conceptual fields* (particularly as concerns the components of a concept: Vergnaud 1990). In particular, in the course at stake the zone of proximal development was evoked several times, when discussions with participants concerned transcripts of classroom interactions in elementary school. Participants tried to identify instances of pupils' argumentative performances beyond the level of their own autonomous elaboration, and how they had been driven by the teacher's prompts or some schoolmates' contributions. Vergnaud's theory served as a tool to identify, in particular, the role of argumentation in the development of pupils' mastery of concepts, in the perspective of Vygotsky's *everyday concepts—scientific concepts* dialectics (see Douek 1999).

Engeström's elaboration on *expansive learning*, related to the development of identity and autonomy in educational programs (Engeström and Sannino 2010), and the construct of *Field of Experience* (Dapueto and Parenti 1999) inspired the planning of the course at stake. In particular, the expansive learning construct served as a reference to engage participants in activities of production of new knowledge on professional relevant issues, and of self-evaluation of their performances related to the aims of the course (see below). The construct of Field of Experience was used to identify suitable cultural domains for argumentation and discuss them with participants, in the perspective of their use for the development of argumentative competencies in elementary school.

The Course

The course on mathematical argumentation was introduced in 2014/15 in the Genoa University curriculum for the master's degree in Sciences of Elementary Education as a fourth year course. Our decision was a consequence of the difficulties met to deal in a homogeneous, coherent way with various aspects of the subject in the other courses of mathematics and mathematics education.

In this chapter, for practical reasons related to available documents, we will focus on the 2015/16 version of the course, similar to that of the previous year.

42 prospective teachers (hereafter called ‘participants’) attended this 24-hour course during a period of 7 weeks, in Autumn 2015 (with at least 12 further homework hours during the course, and more than 36 h of self-reflective activities on the participants’ personal learning trajectory after the end of the course—see below: Participants’ assessment). All their written productions and some audio-recordings of classroom discussions were available. One of the authors was the teacher of the course.

The Design of the Course

The design of the course was inspired by three principles, related to our theoretical framework:

- focus on the CAC of mathematical argumentation, starting from its framing through Toulmin’s model;
- narrow connection, established through the four parts of the course, between analysis of elementary students’ behaviors in classroom argumentative activities, participants’ argumentative activities at an adult level and task design on argumentation (see Zazkis 2008; Stylianides and Ball 2008);
- method choices (including evaluation) in line with the aim of developing CAC competence on mathematical argumentation.

As concerns method and evaluation issues, and focus on the development of CAC competencies, the course is representative of the whole set of mathematics and mathematics education integrated courses of the master’s degree in Sciences of Elementary Education at the Genoa University.

The Content

The course’s main content was mathematical argumentation, for three reasons:

- it is one of the key subjects in the National Guidelines for Curricula of grades I to VIII (6-to-14-year-old students);
- the emphasis on argumentation is new in Italian elementary schools, and no didactical tradition exists at this school level (see Introduction), thus teachers need to be prepared to deal with it in the classroom;
- systematic practice of argumentation is strictly related to the development of mathematical concepts (Vergnaud 1990) when dealing with key content of the elementary school curriculum, in the transition from every day to scientific concepts (Douek 1999).

Identifying mathematical argumentation in students’ mathematical productions from elementary school was the aim of the first part of the course (about 15% of the whole course). Toulmin's model was explicitly introduced to promote this competency, and to prompt reflection on some epistemological and logical-linguistic features of argumentation, when dealing with mathematical content; in particular, the nature of warrants (empirical, general/based on properties, or formal/syntactic).

The aim of the second part (about 40% of the course) was to develop participants' competencies about task design on argumentation, starting with single tasks in different fields of experience (Calendar, Seasons, Money and purchases, Grammar, Classroom productions, Numbers and operations, Space and geometry), then moving to a sequence of tasks on a given content in the mathematical field of experience. Reflection on epistemological and logical-linguistic features of questions, suitable to activate argumentation, helped participants to move from identification of argumentation to the production of tasks in a given context of shared knowledge aimed at developing elementary students' argumentative competencies. In this part of the course, based on participants' task design, analogies and differences of argumentation as a discursive activity across different domains were discussed, with explicit reference to the anthropological relevance of argumentation since the ancient Greeks. Note that in this part of the course, the concern for students' ways of dispelling doubts about the truth of statements (or of falsifying statements), and the focus on different criteria for assessing truth, together with pupils' ways of communicating their reasoning, gradually emerged as a major object of interest for participants, even if Habermas' construct of rationality was not explicitly introduced.

The third part of the course (about 30%) was aimed at identifying different ways of justifying a mathematical statement in elementary theory of numbers, through participants' personal activities of conjecturing and proving in mathematics. Historical (going back to Euclid's arithmetic) and epistemological considerations helped distinguish between algebraic-formal justifications (based on the use of the algebraic language) and verbal-semantic justifications (referring to basic knowledge about numbers) (see Guala and Boero 2017). In this phase, participants became protagonists of argumentation and proving activities, and at the same time evaluators of their mates' productions. The subsequent analysis of 5th grade (10–11 years old) students' productions on the same tasks helped participants to become aware of those students' potential in conducting a deductive reasoning, when inferences are based on the meaning of arithmetic propositions.

The fourth part of the course (about 15%) was aimed at further developing participants' task design competencies. This time, tasks had to be addressed to develop elementary students' mathematical argumentation (general claims and general warrants, related to arithmetic properties known by students).

Teaching Method

The course was organized according to a model of "laboratory course", with 2–3 worksheets to be completed individually each week in the classroom (and 1 or 2 worksheets to be completed each week as homework). Worksheets usually contained an informative part (reminder of the content of previous lectures, and/or brief introduction to new definitions, elements of theory, etc.), followed by open questions. Each worksheet working phase was followed by a discussion, guided by the teacher, based on some participants' productions (selected by the teacher) and/or on some elementary school students' productions on related issues. The role of the teacher was to lead the discussion and to introduce (when appropriate) new cultural elements, theoretical tools, etc. connected to the content of the discussion. The percentage of classroom time devoted to individual work was about 35% (about 30' for each worksheet). Worksheets were collected by the teacher and given back to participants at the end of the course, for self-assessment purposes. Participants' CAC activities were the object of most tasks; thanks to the organization of the course, this choice resulted in participants' production of knowledge on the content of the tasks (cf. Engeström and Sannino 2010), and at the same time it developed their CAC competence. The method of participants' assessment was conceived to create the best conditions for their engagement in CAC activities, during and after the course.

The Participants' Assessment

Participants' CAC competence cannot be assessed through a traditional written or oral exam, because the relationship with the teacher, needed for participants' development of the CAC competence, is incompatible with the kind of subordination to the teacher inherent in the traditional system of evaluation. Wisely formulated interviews might work better than traditional oral exams, but still participants should prepare themselves to answer questions posed by the teacher, not to self-pose questions and develop an inner dialogue on what they have learnt. Usually, participants' relation with mathematics (and other disciplines as well) is dominated by the need of showing to the teacher that they know what the teacher wants them to learn. In the perspective of becoming professionals, participants should move towards an autonomous taking in charge of their knowledge as a corpus of mathematical notions and practices, and reflections about their roots and connections with present culture and children's extra-school experiences. Participants need also to reflect on their own learning trajectory. This is an important occasion to realize how learning (in mathematics, as well as in other domains) is not a linear, ideal path, undermined in reality by possible mistakes, but a very complex, non-linear process, where pitfalls and weaknesses may be important occasions for moving ahead, according to a personal evolution which usually differs from that of ones' own mates. Reflecting on their learning process is also an occasion for participants to learn to analyze their future students' actual learning trajectories out of the traditional right/wrong dichotomy, and to overcome the idea of an 'optimal' universal learning trajectory. As a consequence of the necessity to move towards such an autonomous and active role, in the last years, participants' assessment in our courses has been based on their written retrospective analysis of their personal work during the course and related difficulties, and on their synthesis of their own learning itinerary. In both cases, precise reference is requested to their individual written productions: thus the participants' attention is contractually addressed (during the course, and in the final reflective analytical and synthetic activity) to their growing up as CAC expert teachers. Evaluation criteria by the teacher include the precise correspondence between the performed analysis and synthesis and the participant's written productions, and the quality of her analysis.

Three Snapshots from the Course

We have chosen three short segments from the 2015 course in order to show how planning of the course resulted in teaching and learning activities. Their synthetic presentations ("snapshots") should provide the reader with information on the activities, which have been referred to in the pieces of participants' self-reflective reports to be presented in the next section. The snapshots will also present some typical tasks of the first three parts of the course, and how participants dealt with them. Some tasks have been based on classroom productions of elementary school students, collected by school teachers who performed research activities with us on the development of students' argumentation at the elementary school level.

First Snapshot

This derives from the first part of the course. Participants were requested to identify traces of argumentation (according to their conception of argumentation) in three texts of 2nd grade (7-year-old) students at the end of the year. Those students had to write down a previously learnt criterion to identify even numbers (by checking if the last digit on the right is even), and to justify it according to what had been discussed in the classroom under the guidance of the teacher: each number is a sum of powers of ten and of the number represented by the last digit; if the last digit is even, the given number is even, because it is the sum of even numbers (an already shared knowledge). The first text presents the criterion in a clear, detailed way, without any justification for it:

> I. Even numbers are those integer numbers, which are divisible by 2; odd numbers are those integer numbers, which are not divisible by 2. In order to identify an even number one must check if the last digit is even.

The second text contains not only the criterion, but also a complete justification for it, based on a generic example (Mason and Pimm 1984), expressed in a rather unclear (at a first reading) and not completely explicit way. In our faithful translation into English, we maintained the Italian word "che", used by this student with three different meanings (like in Italian poor speech): as "and" (in the first case), as "because" (in the second and third case), and as a pronoun ("which", in the fourth case):

> II. Even numbers are those divisible by 2, "che" [and] if a number ends by 0, 2, 4, 6, 8 is even, "che" [because] when for instance I take 34 it is even, "che" [because] 30+4 with 30 divisible by 2 and 4 even makes even+even, "che" [which] is even.

The third text clearly presents the criterion and an incomplete justification (the second warrant is lacking). At the end of the worksheet, participants were requested to write down what an argumentation was for them:

> III. Even numbers are identified by the last digit on the right, which must be even. In other words: if the last digit on the right is even, the number is even. Indeed if the last digit on the right is even, we make the addition with the remaining even part of the number.

One half of the participants thought that the first text is argumentative (in line with their conception of argumentation as a "clear presentation of a procedure or an idea"). The other half of participants thought that argumentation consists in providing justifications or motivations for a claim, thus they qualified the first text as non-argumentative, but one third of them did not identify argumentation in the second text, and were in doubt about argumentation in the third text.

After comparison of their productions and discussion, participants were provided with Toulmin's (nuclear) model of argumentation; then they were requested (in worksheet 2) to identify argumentation (and re-write the text accordingly) in the second text, and to complete the argumentation in the third text.

Second Snapshot

This comes from the second part of the course. Participants dealt, in worksheet 7, with the following task proposed in a 2nd grade classroom (7-year-old students):

> Discuss the following text written by your schoolmate Maria:
>
> *Another criterion to check if a number is odd is to add its digits, and check if the sum is odd. For instance 27: 7+2=9 is odd, and 27 is odd.*

Participants had to analyze some children's very rich productions (most participants made an appropriate use of Toulmin's model); then participants were requested to produce other tasks with similar logical characteristics in different domains (arithmetic, grammar, natural sciences, etc.) for 2nd grade and for 4th–5th grade children. One half of the participants met difficulties to produce such kind of tasks. In some cases, they produced false claims, but were not able to find examples that apparently validate them; in other cases they produced true claims with suitable examples to validate them. Some participants' good productions, in mathematics and in other domains, helped their mates to focus on the core element of Maria's text. Here are some sentences proposed by participants for tasks similar to the original one:

> Discuss the following text written by a schoolmate of yours:
>
> (For 2nd grade,) The first names, which end on the right with the letter *a*, identify female people. For instance: Anna ends with *a*, and is the name of a female person.
>
> (For 4th grade.) Odd numbers are prime numbers. For instance: 13 is odd, and is a prime number.
>
> (For 5th grade.) Global warming is a fake. For instance: according to Internet, in Piemonte (the region of Turin), February 2012 was the coldest month of the last century.

The long (55′) discussion about participants' productions (particularly for productions like the third one) brought to the fore the importance of developing children's critical attitude toward this kind of statements: they have been used, and still are used, to convince people about the validity of a statement in different fields. The cultural-anthropological relevance of argumentative competencies emerged in this part of the course as a major motivation for their development since elementary school.

Third Snapshot

This comes from the third part of the course.

Two of the tasks for participants were as follows:

- (Worksheet 17) What can you say about the common divisors of any two consecutive natural numbers? Justify your conjecture in a general way.
- (Worksheet 21) What can be said about the GCD [greatest common divisor] of all the products of two consecutive even numbers? Justify your conjecture in a general way.

Participants identified true conjectures, but one half of them met great difficulties in producing general justifications for them. This happened even in the case of the second task (Worksheet 21), in spite of a detailed discussion of some participants' productions for the first task (Worksheet 17), and of an analysis of 5th graders' productions for the same conjecturing and proving task. Note that participants had experienced proof in high-school for a number of theorems, going from a dozen to more than 40 (according to the type of high-school) presented by the teacher at the blackboard, with reference to the proofs displayed in the textbooks, and then studied at home and individually reproduced to get the passing mark. Reflections on the limitations of transmissive teaching in this domain were one of the follow-ups from this part of the course. The role of generic examples (Mason and Pimm 1984), taken from some children's productions on the same tasks, came again to the fore as a possible transition means from checking the validity of a statement by using examples, to producing a mathematical proof. Also, the difference between algebraic-formal proofs and verbal-sematic proofs is put into evidence, and further reinforced through another conjecture, concerning the GCD of the products of three consecutive natural numbers. In that case, an algebraic-formal proof requires knowledge of combinatorics, or of modular arithmetic; however, a verbal-semantic proof is rather easy by considering the distribution of the multiples of 2 and 3 on the number line. Then some photocopies of pages taken from Euclid's books on arithmetic theorems were distributed and three easy proofs were discussed.

In this part of the course, most participants realized that proving based on semantic inferences was a legitimate way of proving, and that it had been a vehicle of mathematical knowledge development for about two thousand years in Western civilization. The following excerpt from a discussion shows how three participants (P1, P2, P3) contributed to the comparison between algebraic-formal and verbal-semantic proving for arithmetic theorems:

P1 I was surprised when the teacher said that 10-year-old students may prove simple arithmetic theorems. Now I understand that proving arithmetic theorems does not need Algebra (pause) I understand the words are sufficient. 10-year-olds may use words to prove.

P2 Yes, me too, I was surprised with 10-year-olds' proving, because I thought that only proving arithmetic theorems with Algebra was acceptable in Mathematics.

P1 In these photocopies we see a lot of proofs by Euclid, he was able to prove, to establish the truth of those statements.

P3 Only by words and segments! A little bit like in Geometry! Reasoning is based on thinking, on looking at segments and imagining numbers.

P2 As lengths of segments.

Participants also realized that verbal-semantic proving is strictly connected with argumentation in other cultural domains (like the social and natural sciences). In her utterances during a classroom discussion, a participant developed connections between proving in Mathematics and argumentation in Ecology under the teacher's prompts:

P It seems to me that this kind of reasoning is important, is relevant not only in mathematics. (pause) Data, claims and warrants may be different.

Teacher What do you mean by "different"?

P Yes, different (pause). Like in a discourse on warming, on global warming (pause). Data may be (pause) data: increase of CO_2 from cars and industries. Claim may be (pause) expansion of deserts. Warrant: the increase of temperatures as a consequence of increase of CO_2 in the atmosphere.

Teacher What in common with reasoning in arithmetic?

P I don't know exactly (pause) perhaps the logic aspect, the structure of argumentation (pause) perhaps more precisely: verbal reasoning, the use of words according to the same structure. In arithmetic we reflect on the meaning of sentences on numbers, (pause) while here there are more complex things.

(The teacher writes the proposed example concerning global warming on the blackboard, according to Toulmin's model.)

P Yes, I see a claim, in this case a real fact, supported by a warrant, in this case one of the effects of the increase of CO_2 in the atmosphere.

The two reported excerpts provide evidence about how CAC competence has been developed during the course. However we may observe how in both cases the lack of a suitable tool (which might have been the rationality construct) did not allow participants to deepen the comparison between different kinds of proving (in the first case); and to clarify the nature of the common aspects and the differences between argumentation in the two domains, in the second case. We will discuss this issue in the last section.

Analysis of Participants' Self-reflective Reports

The evaluation of the effectiveness of the choices performed to frame, plan and manage the course concerns different aims (including the development of task-designing competencies and the use of Toulmin's model for argumentation). Here, we consider the development of CAC competencies on the subject of mathematical argumentation. Criteria, chosen to evidence the quality of participants' CAC specific performances in self-reflective reports, are related to the nature of the CAC competence and to Habermas' construct, as adapted to to mathematical argumentation:

(i) the identification and precise wording of the nature of argumentation (in terms of the kind of warrants and strategic choices: semantic or syntactic warrants; a list of examples, or generic examples, or strategies aimed at generality in the reasoning; etc.) in her own, or other participants', or children's texts;
(ii) the identification and precise wording of lacking warrants, or illegitimate inferences, or bad strategic choices; in her own, or other participants', or children's mathematical productions;
(iii) the identification (through CAC when appropriate) and precise wording of difficulties (and their roots) in the participant's learning trajectory, be they overcome or not..

The following quotes from participants' final self-reflective reports are related to the three reported snapshots (e.g. A2 refers to the second snapshot). Superscripts [i), ii), iii)] will be inserted in the quotes in order to put into evidence some instances of our coding of participants' texts according to the criteria (i), (ii), and (iii) we described earlier.

Participant A is a representative of the 30% of higher level participants—those whose final self-reflective reports systematically satisfied all the above criteria.

A1: In worksheet 1, I was not able to identify the second text as the only one being fully argumentative, because my idea of argumentation was influenced by a model of clear presentation of one's ideas, even if no justification for them was included [iii)]. The concatenation of two Toulmin's cycles was another problem met by me in the second worksheet [iii)]. I tried to get a unique step of reasoning, with an only partial warrant for it [ii)]. Through comparison with my colleagues' productions, I started to identify the usefulness of Toulmin's model as a tool for analyzing argumentation, particularly in complex argumentation, when the claim of the first step becomes a data for the second step (and so on) [i)].

A2: While I was able, in worksheet 7, to produce other tasks in mathematics similar to that based on Maria's idea, I met a lot of difficulties to produce similar tasks in other non-mathematical domains. It was like if a wall separated mathematics from other domains [partial iii)]. My task: "Find the mistake in the sentence: *the water freezes because the temperature is* –5 °C" did not correspond to Maria's task, because the warrant is a case of a more general correct justification [ii)]. I did not understand, at the beginning, the common structure of a local coincidence with no general correspondence in the relationship between warrant and claim [i), iii)]. Some tasks produced by my colleagues (….) helped me to both reach deeper understanding about the nature of Maria's mistake, and to identify the importance for life of the competence at stake.

A3: I engaged a lot in proving tasks from worksheets 17 and 21 by using algebraic methods, because I thought that only in that way were a proof a true mathematical proof [iii)]. I was unsuccessful in both cases, because I was not able, after more than 3 years, to use the algebraic language (learnt in high-school) [iii)]. In the first case, I was not able to get any useful algebraic formalization; in the second case, I made a mistake by writing $2(n+2)$ instead of $2n+2$ to represent the second even number [ii)]! Then I was surprised by two facts: the possibility of a non-algebraic deductive proof based on knowledge of elementary arithmetic properties (as a legitimate mathematical proof, since Euclid) [i)], and the fact that such proof was accessible to 5th grade students, against my idea about their limited reasoning skills (probably related to my idea of mathematical reasoning as essentially based on manipulation of algebraic symbols [iii)]). The use of a generic example for the task of worksheet 17 was difficult to accept for me because it was related to the use of examples (in high-school I learnt that examples may be used only to disprove conjectures [iii)]). But now I realize that if we consider the example of 15 and 16 and we see that 5 (a divisor of 15) cannot be a divisor of 16 because the remainder is 1, and the same happens with 3, then this way of reasoning may be generalized to get a general deductive proof (based on 1 as the remainder of the division of the second number by a divisor of the previous number, different from 1) [i)].

Participant B is a representative of the group of 30% lower level participants—those whose self-reflective reports only partially satisfied, here and there, some of the above criteria:

B1: While dealing with worksheet 1, I had a vague idea of argumentation as a logical matter related to inferential reasoning (according to my memory of what I learnt about Aristotle in the high-school in the course of Philosophy), but I did not succeed in identifying argumentation in the second text and in re-writing it as an argumentative text (worksheet 2) [partial ii)]. I was not able to use Toulmin's model at that time because it was something very abstract and far from actual texts [iii)]. Now I would be able to use it, as I did in the last worksheet. And now I think to better understand what Aristotle wrote.

B2: It was (and it still is) difficult for me to find the deep reason why Maria's criterion is not acceptable. Yes, I realized that it did not work in the case of 35, but what was the underlying logical pitfall? Some of my colleagues explained it in different ways, but still the core idea is not clear for me [partial iii)]. After some difficulties, I produced two similar tasks in other domains in Worksheet 7, but I must say that it was by imitation, not by understanding [partial iii)].

> B3: I was able to validate the statement of worksheet 17 by using algebraic language [i)] (I was a good student in mathematics in high-school), but not the statement of worksheet 21. Concerning alternative ways of proving it, by reference to the properties of the number line, I realized that it is possible to prove the second statement by making reference to that knowledge [partial i)]. But in the case of the first statement, the use of a generic example for me does not escape the reference to numerical examples, thus: why may we consider it as a bridge towards a GENERAL proof? [partial iii)]

Participants' performances in their self-reflective reports provide some evidence about the maturation of the CAC competence in the case of mathematical argumentation not only among the highest level participants, but for all of them, according to the chosen criteria (at least as concerns, for some issues, the awareness of personal persistent difficulties and not achieved aims—see B2, B3). In the reported excerpts, we may identify phenomena occurring in many other reports: how previous historical and epistemological knowledge became a component (integrated with further knowledge) of the maturation of the CAC competence (A3, B1); and how participants were enabled to put into question previous conceptions about argumentation (A1) and proof (A3), and to identify crucial epistemological (A1) and anthropological (A2) aspects of argumentative competencies, even if in some cases maturation did not result yet in full mastery of them (B2, B3).

Discussion

Some elements of the self-reflective reports suggest that a crucial role for participants' maturation of their CAC competence was played by the evaluation system *integrated* with the methodological choices of the course. The participants' practice of CAC was guided through the worksheets, then collectively shared and enhanced during the related discussions. After the end of the course a precise correspondence between participants' self-evaluation and personal worksheets was required. All this allowed participants to identify their strengths and weaknesses concerning CAC not only for their teacher but also *for themselves*, thus opening the way to the autonomous practice and improvement of CAC. Participants moved from specific *knowledge* on argumentation and proof, acquired during the course (or before), towards *its conscious use*: to analyze cultural achievements; to establish links between different cultural domains (see A2); and to identify different, valuable mathematical practices related to historical and personal evolution of mathematical reasoning (see A3). All of them are important components of teachers' professional competence related to CAC.

Looking back to the course through the transcripts (cf. the reported snapshots), we have seen how participants gradually became accustomed to being responsible for their knowledge. Questions traditionally addressed to teachers in school (and even at the university level, in standard mathematicscourses or education sciences courses), like "Is what I wrote correct?", gradually became self-addressed questions —a fundamental step towards learning to evaluate textbooks and analyzing students' productions in the school in an autonomous way. But the analysis of participants' self-reflective reports, and the comparison with participants' performances in parallel courses for secondary school teachers (see Guala and Boero 2017), raises a question: what about the explicit introduction in this course (or in another course) of Habermas' construct of rationality? Indeed participants lack a vocabulary (and a perspective) to identify different rationalities (within mathematics, and in the comparison with other cultural domains), which might contribute to their CAC; and also to get a deeper insight into their own work and into students' work in terms of epistemic, teleological and communicative components. Constraints resulting from limited available time and participants' epistemological and historical background might be overcome through a different coordination of content with the other courses.

What happens when our prospective teachers enter the school? We may say that sporadic follow-ups at schools, after the degree, of participants in our courses provide evidence about the difficulties met by some of them in using their CAC competence. Difficulties are mainly due to the fact that educational choices resulting from CAC are scarcely compatible with both the transmissive teaching that prevails in schools and the present school culture (especially when teaching is planned by teams of teachers, as is usual in many Italian elementary schools) (cf. Stylianides et al. 2013, for a similar situation in another country). However, it is interesting to observe that positive results are emerging in schools (usually situated in "problematic" districts) where several young teachers coming from the present university teacher education program may work together. We have kept in touch with some of these young teachers. Their requests for suggestions from us and their examples of performed teaching activities on argumentation show how they try to develop argumentative skills by identifying suitable tasks in different cultural domains, according to the opportunities offered by local situations. They also engage in analyzing students' performances in those activities beyond the "right-wrong" dichotomy, by trying to identify the features of the students' individual argumentative processes.

References

Arzarello, F., & Bartolini Bussi, M. (1998). Italian trends in research in mathematics education. In J. Kilpatrick & A. Sierpinska (Eds.), *Mathematics education as a research domain* (Vol. 2, pp. 243–262). Boston: Kluwer.

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

Boero, P. (2006). Habermas' theory of rationality as a comprehensive frame for conjecturing and proving in school. In *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 185–192). Prague: PME.

Boero, P., & Guala, E. (2008). Development of mathematical knowledge and beliefs of teachers: The role of cultural analysis of the content to be taught. In P. Sullivan & T. Wood (Eds.), *The international handbook of mathematics teacher education* (Vol. 1, pp. 223–244). Rotterdam: Sense Publishers.

Boero, P., Guala, E., & Morselli, F. (2014). Perspectives on the use of the Habermas' construct in teacher education: Task design for the cultural analysis of the content to be taught. In *Proceedings of the 38th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 228–232). Vancouver (CA): PME.

Boero, P., Douek, N., Morselli, F., & Pedemonte, B. (2010). Argumentation and proof: A contribution to theoretical perspectives and their classroom implementation. In *Proceedings of the 34th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 179–205). Belo Horizonte: CEDECOM, UFMG.

Dapueto, C., & Parenti, L. (1999). Contributions and obstacles of contexts in the development of mathematical knowledge. *Educational Studies in Mathematics, 39,* 1–21.

Douek, N. (1999). Argumentation and conceptualization in context: A case study on Sun shadows in elementary school. *Educational Studies in Mathematics, 39,* 89–110.

Durand-Guerrier, V., Boero, P., Douek, N., Epp, S., & Tanguay, D. (2012). Examining the role of logic in teaching proof. In G. Hanna & M. De Villiers (Eds.), *Proof and proving in mathematics education* (pp. 369–389). New York: Springer.

Engeström, Y., & Sannino, A. (2010). Studies of expansive learning: Foundations, findings and future challenges. *Educational Research Review, 5,* 1–24.

Grabiner, J. V. (2012). Why proof? A historian's perspective. In G. Hanna & M. de Villiers (Eds.), *Proof and proving in mathematics education* (pp. 147–167). New York, NY: Springer.

Guala, E., & Boero, P. (2017). Cultural analysis of mathematical content in teacher education: The case of elementary arithmetic theorems. *Educational Studies in Mathematics, 96*, 207–227.

Habermas, J. (1998). *On the pragmatics of communication*. Cambridge, MA: MIT Press.

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

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

MIUR. (2012). *Indicazioni Nazionali per il Curricolo della scuola dell'infanzia e del primo ciclo d'istruzione*. http://hubmiur.pubblica.istruzione.it/alfresco/d/d/workspace/SpacesStore/162992ea-6860-4ac3-a9c5-691625c00aaf/prot5559_12_all1_indicazioni_nazionali.pdf.

Pellerey, M. (1989). *Oltre gli insiemi*. Napoli (It): Tecnodid.

Radford, L., Boero, P., & Vasco, C. (2000). Epistemological assumptions framing interpretations of students understanding of mathematics. In J. Fauvel & J. van Maanen (Eds.), *History in Mathematics Education*. *The ICMI Study* (pp. 162–167). Dordrecht (NL): Kluwer.

Schoenfeld, A. H. (1994). Reflections on doing and teaching mathematics. In A. H. Schoenfeld (Ed.), *Mathematical thinking and problem solving* (pp. 53–75). Hillsdale, NJ: Lawrence Erlbaum Ass.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4–14.

Siu, M. K. (2012). Proof in the western and eastern traditions: Implications for mathematics education. In G. Hanna & M. de Villiers (Eds.), *Proof and proving in mathematics education* (pp. 431–442). New York, NY: Springer.

Stylianides, A., & 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*(4), 307–332.

Stylianides, G., Stylianides, A., & Shilling-Traina, L. (2013). Prospective teachers' challenges in teaching reasoning-and-proving. *International Journal of Science and Mathematics Education, 11*(6), 1463–1490.

Toulmin, S. E. (2008). *The uses of argument* (8th ed., 1st English ed. in 1958). Cambridge: Cambridge University Press.

Tsamir, P. (2008). Using theories as tools in mathematics teacher education. In D. Tirosh & T. Wood (Eds.), *The international handbook of mathematics teacher education* (Vol. 2, pp. 211–234). Rotterdam: Sense Publishers.

Vergnaud, G. (1990). La théorie des champs conceptuels. *Recherches en Didactique des Mathématiques, 10,* 133–170.

Zazkis, R. (2008). Examples as tools in mathematics teacher education. In D. Tirosh & T. Wood (Eds.), *The international handbook of mathematics teacher education* (Vol. 2, pp. 135–156). Rotterdam: Sense Publishers.

Chapter 5
Toward an Evolving Theory of Mathematical Practice Informing Pedagogy: What Standards for this Research Paradigm Should We Adopt?

Keith Weber and Paul Christian Dawkins

Abstract In this chapter, we provide commentary on the four preceding chapters on proof in mathematics education. We contend that each of these chapters considers how Mathematical Practice can inform Pedagogy (MPP) research. We use these chapters to begin a discussion on what factors mathematics educators should consider when producing and evaluating MPP research. Each chapter seeks to inform mathematics education using the philosophy or history of mathematics. We argue that our field continues to borrow from these relevant fields without clear criteria for evaluating such research and without a framework for the transposition across disciplines. The chapters also all entail meta-mathematical learning goals for students and pre-service teachers. We raise questions about the exact intent of these learning goals and assessment of such learning whose answers would enhance the contributions of MPP research.

Keywords Mathematical practice · Mathematics · Pedagogy · Proof Teaching

Mathematical Practice Informing Pedagogy

This chapter represents the authors' reflections on and response to the previous four chapters. In attempting to conceptualize the four chapters separately and as a whole, we formulated a description of what (in our understanding) they shared with one another and with some prior literature. We call this common pattern of research "mathematics practice informing pedagogy," which we explain and explore hereafter.

K. Weber (✉)
Rutgers University, New Brunswick, USA
e-mail: keith.weber@gse.rutgers.edu

P. C. Dawkins
Northern Illinois University, DeKalb, USA
e-mail: pdawkins@niu.edu

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_5

Commonalities in These Book Chapters

In contemporary mathematics education, we generally accept that there is a relationship between what professional mathematicians do and how mathematics is taught. This is especially the case with mathematical proof (Dawkins and Weber 2017; Weber et al. 2014). In particular, the four chapters in this volume illustrate that the practices of mathematicians can inform what it is that we want students to learn and how instruction should be designed. We observed two commonalities among these chapters. First, each chapter is consistent with Manin's famous dictum that "good proofs are proofs that make us wiser" (cited in Aigner and Schmidt 1998). In their own way, each chapter illustrated how the role of proof in the classroom goes beyond persuading students that mathematical claims are true, but advances some other pedagogical goals as well. Second, each chapter uses mathematical practice to inform what those pedagogical goals ought to be and how those pedagogical goals might be achieved. For rhetorical convenience, we will call research that satisfies the second commonality as Mathematical Practice informing Pedagogy (MPP) research. MPP research on proof has been present in mathematics education for some time, at least since De Villiers' (1990) seminal paper on the purposes of proof, and several of the authors of these chapters previously have made important contributions to MPP research (e.g., Boero 2007; Durand-Guerrier 2016; Hanna 1990; Harel 2007). We believe that the authors' chapters represent a promising trend toward MPP research in the mathematics education literature on the teaching and learning of proof.

The aim of this chapter is to use the four chapters to critically analyze MPP research. While MPP research on proof is becoming increasingly common, the field currently lacks standards that MPP research should meet and criteria by which it can be evaluated. In this section, we summarize the contributions of each chapter. In our summaries, we highlight how each chapter satisfies the two commonalities that we discuss in the previous paragraph, describe how we think each chapter significantly advances the field, and raise critical questions that came to us as we read the chapter. In the next section, we attempt to synthesize the critical questions that we listed. In particular, we list broad criteria that we think the field ought to consider when evaluating MPP research and we will highlight the challenges for finding systematic ways to form explicit criteria.

Summaries of the Contributed Chapters

Hanna: Criteria for Mathematical Explanations in the Classroom

The goal of Hanna's chapter was to explore what makes an explanatory proof in a pedagogical setting. Hanna's premise was that in mathematical practice, proofs often provided mathematicians with something beyond conviction that a claim was true; proofs sometimes provided mathematicians with an explanation for why claims were true. Hence, Hanna was using mathematicians' practice as a basis for what our pedagogical goals with respect to proof should be. In the past, other researchers have also claimed that pedagogical proofs should be explanatory when possible (e.g., De Villiers 1990; Hanna 1990; Hersh 1993) but there is a lack of agreement on what explanatory pedagogical proofs should be (c.f., Weber 2010a). Hanna noted that we cannot simply adapt philosophers' definitions of explanatory proofs because philosophers and mathematics educators have different disciplinary aims. However, she argued that we can form analogs for some of philosophers' characterizations that are appropriate for mathematics education purposes.

Hanna distinguished between intra-mathematical explanations (explanatory proofs in philosophy) and pedagogical explanations as the two operate within different contexts and assume vastly different levels of expertise. Inasmuch as there is precedent for translating such philosophical criteria to pedagogical contexts (even in our own work: Dawkins 2015), it is necessary to address the translation explicitly. The three criteria for a proof being explanatory identified (regarding intra-mathematical proofs) are proofs depending upon a characteristic property, proofs that support theoretical unification, and proofs that use salient properties. Hanna provides examples of proofs that could be used pedagogically that, in her estimation, satisfy each of these criteria. The goal, however, is that these criteria could serve a generative and/or explanatory role in talking about when and why a proof is pedagogically explanatory.

Hanna's chapter makes several important contributions. Her criteria for a proof being explanatory that could guide the development of pedagogical examples of proof or proving activities. More generally, Hanna raised the important point that the translation from philosophers' interpretation of mathematical practice to pedagogical recommendations is complicated, particularly because different stakeholders (philosophers, mathematicians, researchers, and teachers) have different aims. Hanna's chapter can serve as a model for how an MPP researcher can manage this complexity in making practical suggestions that satisfy the needs of all the different stakeholders.

Reading the chapter raised the following questions to us. To what extent does mathematical explanation actually guide mathematical practice? What evidence would be appropriate to persuade mathematics educators of the importance of explanation in mathematical practice? How can we assess when students perceive characteristic properties, theoretical unification, or salient properties and the influence of such judgments? Since even university mathematics students cannot distinguish between a proof and an invalid argument (e.g., Ko and Knuth 2013; Selden and Selden 2003; Weber 2010b), is it feasible for proof to provide explanation for all students? Would a focus on explanation take emphasis away from the more urgent goal of developing students' ability to gain conviction from proofs and follow the logical chain of reasoning within a proof?

Durand-Guerrier and Tanguay: Systematizing the Real Line

Durand-Guerrier and Tanguay's chapter highlighted several ways in which mathematicians have systematized the real line. Durand-Guerrier and Tanguay observed that the role of definitions and axioms not only affects what needs to be justified and what is permissible in a justification—a point Mariotti (2006) previously highlighted with regard to reference theories—but the authors also noted that different systematizations emphasize and downplay different aspects of the real line (e.g., order relations, completing the rationals, decimal representations). Dedekind cuts accentuate the linear ordering properties of the real line while minimizing the metrical properties of the real line while Cauchy sequences do the opposite. The authors highlighted how studying the proofs in these (and other) systematizations provides learning opportunities for students, but the learning opportunities that are present depend upon the systematization being used. Durand-Guerrier and Tanguay used the different systematizations that mathematicians generated for the real line to show how different proofs (within these systematizations) could contribute to students' understanding of the real line.

This chapter makes several important contributions: The foundational properties of the real line are important for students to distinguish and coordinate, which might be fruitfully done via various systematizations or even by comparing different systematizations. Further, this chapter highlights that the systematizations that we choose can have conceptual consequences for students based on what the systematization takes for granted and what it constructs. The proofs within these systematizations are not merely bookkeeping but can be an important pedagogical device toward understanding the relationship between properties of a concept.

We considered the following questions as we read this chapter: Do the authors' intended pedagogical goals include merely the properties of the real number line or do they also include aspects of mathematical axiomatization? If understanding axiomatization were a meta-goal of instruction, how would we know if it were achieved (a question with which we have wrestled in our own work, Dawkins 2017, in press)? How could we engage students with one or more of these axiomatizations to productively challenge students' implicit realism about the number line (meaning students sense that numbers simply exist rather than being constructed or defined)?

Harel: Solving Cubics to Produce an Epistemic Justification for Complex Numbers

The goal of Harel's chapter was to consider how to produce a student's intellectual need for complex numbers. He presented a historical analysis of how complex numbers emerged from a proof that the cubic equation could solve cubic polynomials and he argued that this proof can create a similar epistemological justification for complex numbers on the part of the student. Thus, Harel was using mathematicians' practice to inform how proof can contribute to the pedagogical goal of establishing intellectual need for new mathematical objects.

In his chapter, Harel used mathematicians' justification for the cubic equation as a locus for postulating various forms of epistemological justification for mathematical knowledge, meaning an individual's conscious link between knowledge and the problem resolved by that piece of knowledge. He proposed a sequence of tasks reflecting historical developments that contributed to the legitimacy of complex numbers within mathematical practice. Harel noted that secondary teachers may lack understanding of why mathematicians chose to create a solution to a problem such as "What is the solution to $x^2 = -1$?" (as opposed to "What is the solution to $x + 1 = x$?"). This is one instance of the broader mathematical practice of expanding mathematical systems to include formerly nonsensical entities (complex numbers, non-Euclidean geometries, intersections of parallel lines, infinity as a number or infinite cardinalities). Harel traced Cardano's method of solving cubic equations, which yielded a complicated set of formulas that, even when resolving in integer solutions, require computation with complex numbers. Harel posited that this counterintuitive phenomenon could warrant legitimizing complex numbers for teachers in a similar way that it did for Cardano and his contemporaries. The underlying principle is to demonstrate the coherence and consistency of operating with complex numbers to solve this accessible task in which they unexpectedly arose. This then provides the grounds on which Harel postulated three types of epistemic justification. Part of Harel's argument for this endeavor and the distinctions among the epistemic justifications is the anticipated value for K-12 teachers to hold stronger epistemic justifications for the mathematics they teach.

Harel's chapter makes several important contributions. His task sequence can be used by practitioners to develop better epistemic justifications to motivate the complex numbers. More broadly, Harel illustrated the general theme that proofs can be used to motivate an epistemic need, a common phenomenon in mathematical practice that is usually absent in mathematics classrooms.

As we read this chapter, we wondered about the following. Harel's analysis was informed by the historical development of the cubic equation, but the historical development was only discussed briefly and important details such as the competitive mathematical environment at the time (creating the *social need* to solve equations that your mathematical opponents could not) were omitted. Harel acknowledged that his historical treatment was not comprehensive, saying that this aspect of the paper would be only "discussed fleetingly". Would a more extended treatment of the history of mathematics have strengthened the chapter or would it be extraneous? How important is epistemic justification for the instruction of complex numbers within the secondary curriculum compared with the other forms of knowledge we might encourage teachers to construct? How can we measure the existence and influence of such meta-conceptions among practicing teachers?

Boero, Fenaroli, and Guala: Contextualizing and Expanding Elementary Preservice Teachers' Conceptions of Proof

Whereas the other chapters presented broad epistemological ideas for how we can inform mathematical instruction, Boero, Fenaroli, and Guala's contribution differed in that it contains an empirical study with evidence of student learning. The study reported on the nature and efficacy of a curriculum that the authors developed to foster particular understandings of proof. The authors call the desired competence Cultural Analysis of the Content (CAC). This reflects their goal that elementary teachers recognize and legitimize forms of argumentation that were acceptable at earlier points in history within mathematical practice (even though some of these forms might not be acceptable to contemporary mathematicians) and are appropriate for elementary instruction: most notably arguments by generic example. Like the other chapters, here we see the authors using some aspect of mathematical practice (a wide variety of proofs have been permissible in different historical periods) to inform the goals of pedagogy (future teachers should be aware that there are many kinds of proofs and open to different types of students' justifications).

The cycle of mathematical activities and reflection on student arguments strikes us as rich and useful. The authors report meaningful shifts in student conceptions of acceptable proof as measured by their reflective writings. Specifically, several participants improved their ability to recognize generic examples and appreciate their implicit generality. Some students fell short of some target understandings in part because they struggled to recognize when two arguments are of the same form (an implicit test for validity). This finding is certainly important and warrants further exploration. This chapter portrays a nice model for fostering meta-conceptions about proving and trying to assess their development in situ.

This chapter motivated the following questions: Given that the proof standards assumed in the project are less compatible with contemporary mathematicians' views of proving, how do the goals of this project translate to secondary and tertiary mathematics instruction? To what extent are *modern* views of proof the aim of elementary education and to what extent should elementary education privilege the standards of today's mathematicians? To what extent should *earlier* views of proving that are no longer held by contemporary mathematicians inform instruction in elementary mathematics teaching? Are the particular methods of proving championed in this approach as valuable for the instruction of secondary and university students?

Framing the Authors' Contributions in a Broader Paradigm

The Need for a Broader Paradigm

We argued that the four chapters that we summarized were instances of MPP research. In each case, the authors of these chapters highlighted some purpose that proof served for the mathematical community. The authors then recommended that pedagogical proofs can serve the same role in a classroom community. The authors then advocated providing students with examples of justifications that satisfied these criteria to provide students with clear vistas for these high-level ideas.

As mathematics educators continue to produce MPP research, we contend that shared criteria for evaluating this research are needed. At present, we observe that MPP research has developed in a piecemeal manner without much consistency in how authors support their arguments. We posit that our research community lacks clear standards for deciding whether particular MPP research makes a meaningful contribution to mathematics education research or whether the pedagogical recommendations developed in this research are promising. As a result, we personally find that we largely evaluate the quality of MPP contributions by the extent to which they resonate with our way of thinking. Of course, this is not a viable way to develop a robust and coherent literature corpus, especially as mathematics educators notoriously disagree about the nature of proof itself (e.g., Balacheff 2008). We elaborate on this point throughout this section.

In this section, we introduce three broad issues that we think mathematics educators should consider when evaluating MPP research. These issues are:

(i) By what standards should we judge mathematics educators' assertions about mathematical practice?
(ii) If we form pedagogical goals based on (particular views of) mathematical practice, how can we determine if these pedagogical goals have been achieved?
(iii) Are the pedagogical goals endorsed by mathematics educators sufficiently important to be included in an already overcrowded mathematics curriculum?

In other words, how do we verify that any stated pedagogical goals are true to mathematical practice, available for assessment, and essential for instruction? We contend that these issues are critical for evaluating MPP research but generally have received scant attention in the mathematics education literature. In particular, these issues were not discussed in the chapters that we are commenting upon nor are they discussed in our own MPP work.

Evaluating Claims About Mathematical Practice

If researchers are going to make pedagogical recommendations based on claims about mathematical practice, then it is desirable that these claims accurately reflect how mathematicians practice their craft. Hence, it is reasonable to ask the following: How accurate do claims about mathematical practice need to be? Should the practices attributed to mathematicians reflect the views or behaviors of the majority of mathematicians (measured by some chosen threshold) or merely be acceptable within identifiable communities of mathematical practice? To what extent is it incumbent upon the authors to justify their claims about mathematical practice? How should mathematics educators evaluate the justifications that are provided? The answers to these questions are not straightforward. Claims about mathematical practice in MPP research often utilize findings and techniques from traditions outside of mathematics education. The authors in this volume borrowed from the philosophy and the history of mathematics; other MPP researchers have adapted methods and cited findings from disciplines such as psychology and sociology. The authors of MPP research generally do not have formal training in these disciplines; the reviewers and readers of MPP manuscripts usually lack the expertise to rigorously evaluate research from these disciplines. Even if the authors and readers shared the background to evaluate the claims about mathematical practice in an MPP paper, it still might be undesirable for authors to comprehensively warrant their claims about mathematical practice. If the authors provided extensive justifications for their claims about mathematical practice, this could draw attention away from the pedagogical recommendations that are the main point of the paper and lead to papers that violate the word count limits in some journals.

What we have observed is that MPP researchers frequently warrant their claims about mathematical practice by appealing to select mathematicians and commentators whose stances align with contemporary mathematics educators while ignoring other mathematicians and commentators who express alternative points of view. For instance, the distinguished mathematician William Thurston wrote an essay in which he argued that mathematical knowledge is contained in the mental models of individual mathematicians and the social fabric of the mathematical community rather than the formal proofs contained in mathematics papers (Thurston 1994). Many mathematics educators who conduct MPP research, including us, have cited Thurston's essay as evidence that informal mathematical reasoning and social community are more important than logical formalism for the growth and verification of mathematical knowledge. However, Arthur Jaffe and Frank Quinn, eminent mathematicians in their own right, responded to Thurston's essay. Jaffe and Quinn (1994) objected that Thurston's perspective was not representative of the mathematical community as the broader community mostly viewed the formal proofs that appeared in papers as the primary sources of mathematical knowledge. Jaffe and Quinn are seldom cited by mathematics educators. Similarly, mathematics educators are keen to cite Lakatos' (1976) *Proofs and Refutations* to stress the fallibility and corrigibility of mathematical knowledge, yet rarely acknowledge well-known philosophical critiques of Lakatos' work (e.g., Feferman 1998). As mathematicians are a heterogeneous group who hold a variety of viewpoints, there is a wide range of conflicting claims about mathematical practice held by individual mathematicians and commentators. Consequently, one can use the quotation of an individual mathematician to support a wide range of inconsistent claims about mathematical practice.

Let us consider these ideas in the context of Hanna's chapter. Hanna makes an empirical descriptive assertion about mathematical practice:

> [A]lmost all mathematicians make the very useful distinction between proofs that only demonstrate that a fact is true and proofs that also show why it is true. The latter are known as 'proofs that explain'.

Hanna does not justify this assertion other than to say that "philosophers have come to a consensus on the central importance of explanation",[1] but Hanna uses this claim as a point of departure to illustrate ways that explanatory proofs can play a useful role in mathematical classrooms.

It is fair to ask, is Hanna's claim true? Do *almost all* mathematicians distinguish between proofs that explain and proofs that do not? Several scholars have observed that mathematicians rarely discuss mathematical explanations (Avigad 2006; Resnik and Kushner 1987; Zelcer 2013). Zelcer (2013) extended this argument, noting that references to mathematical explanation are sparse even among mathematicians who write reflective essays on their craft. Zelcer reasoned that "if mathematics did countenance explanations then we would expect to find more discussion by mathematicians. But we find almost no such discussion" (p. 180). Mejia-Ramos and Inglis (2017) corroborated Zelcer's claims. Mejia-Ramos and Inglis first observed that the instances of mathematicians discussing explanation in the philosophical literature are largely attributed to three individual mathematicians —Poincare, Halmos, and Thurston. Mejia-Ramos and Inglis then reported the results of a corpus analysis, finding references to proofs that "explain why" occur infrequently in the mathematics literature. As a point of comparison, Mejia-Ramos and Inglis observed that talk of explanations for why phenomena occur are far more common in the physics literature. We do not raise these issues to say that Hanna's assertion about mathematical practice is wrong, only that this assertion hardly represents settled science.

We use our preceding analysis to raise several questions about Hanna's chapter that have relevance for all MPP work.

(i) If an MPP author makes an empirical claim about mathematical practice, to what extent is she obligated to support it? Should empirical claims about mathematical practice require empirical evidence? Is it appropriate to justify

[1]Here, it might be better to state, "philosophers *who study mathematical explanation* have come to a consensus that mathematical explanation is important". In Horsten's (2016) survey of the philosophy of mathematics, mathematical explanation is only cursorily mentioned as one of several future topics that are generating interest. There might not yet be enough interest in the philosophical community for a consensus to be formed on this topic.

claims about almost all mathematicians by appealing to individual mathematicians who might not be representative of their communities?

(ii) Claims about mathematical practice, or more generally the philosophy or history of mathematics, are rarely "settled science." To what extent should MPP researchers discuss debate around these issues? Hanna acknowledged that skeptics of mathematical explanation such as Zelcer existed, but quickly dismissed Zelcer's position "as a minority view" in the philosophy community, which is (to our knowledge) an accurate assertion. Is this sufficient? More generally, would delineating the contours of the debate about the philosophy or history of mathematics be illuminating to the reader or would it merely waste valuable journal space and distract the reader from the more important points that the author was trying to make?

(iii) To what extent does the accuracy of MPP researchers' claims about mathematical practice matter? In Hanna's case, suppose that only *some* mathematicians distinguished between explanatory and non-explanatory proofs. Or even suppose that only *a few* mathematicians did so. To what extent would that weaken the support for the pedagogical recommendations that she provided? It may be that the most honest justification for some aspects of mathematics education practice are located in the transposition of mathematical proof to the classroom rather than in mathematicians' practices of proof. In this case, it would be helpful to accurately locate the warrants for these beliefs and practices.

To avoid misinterpretation, the questions we raised about Hanna's chapter should not be read as critiques as her work. Her contribution was in accord with the current norms for MPP research. Instead our intention is to open about a more general conversation about how MPP research should be presented. If Hanna has committed any sins of omission in her chapter, we have been guilty of the same transgressions in our own MPP work. Indeed, the issues that we raised are pertinent to all the epistemological chapters in this volume. Regarding Durand-Guerrier and Tanguay's chapter, the authors exposit various axiomatizations of the real line without precisely laying forth their interpretation of the role and nature of axioms within mathematical practice. Some constructions begin with the limitations of the rationals while others assume the decimal representation of numbers. Some constructions extend naturally to the plane while others do not. By what standard to we evaluate these options for the classroom? In our own work on axiomatizing in planar geometry (Dawkins 2017, in press), we observe the recurrent alternation between students interpreting axioms as *referential descriptions* of familiar objects (e.g., the real line, the Euclidean plane) and as *stipulated constructions* of mathematical objects (e.g., an irrational number, the Hyperbolic plane). Traditional views of mathematical practice hold that axioms, like definitions, are stipulated rather than referential, which suggests that students should view them as such. However, from a cognitive standpoint axioms of familiar objects like the real line are referential in creation and stipulated in application. Dawkins (in press) observes both limiting and productive aspects of referential reasoning about axioms. Durand-Guerrier and

Tanguay do not convey their view of axiomatizing in mathematical practice nor their goals for student interpretation of constructions of familiar objects (like irrational numbers). We see this as another instance of how MPP research would benefit from a clearer framework for characterizing mathematical practice and how to transpose those practices to the classroom at various levels of instruction.

Assessing Pedagogical Goals

Each of these chapters set forth high-level goals for student learning—what we consider meta-conceptions of mathematical practice. We recognize that this represents part of the broader shift in mathematics education toward engaging students in authentic mathematical processes so they can adopt (and teach) mathematical epistemologies. Being that we have published studies in a similar vein (e.g., Dawkins 2017), we recognize the value in such work and appreciate the inherent difficulty in assessing such knowledge. Nevertheless, we observe that only Boero et al. provide clear means for assessing the targeted meta-conceptions. To avoid misinterpretation, this is not a critique of these chapters; assessments should only be designed after researchers have a good theory of what it is that should be assessed. Further, lack of clear means of assessing knowledge does not invalidate efforts to support student learning of that knowledge.

We believe this issue is of paramount importance if the mathematics education community intends to build on these chapters and design instruction to achieve the pedagogical goals that the authors proposed. It would be unreasonable for us to expect the authors to propose specific assessments in their chapters. What we suggest instead is that MPP researchers can provide *directions* for assessment. Let us consider Harel's chapter on sequencing proofs to develop the epistemic need for complex numbers. How would a researcher or a teacher determine if Harel's sequence, or another instructional sequence, was successful in achieving Harel's goals? Would this involve a qualitative analysis of individual students' responses to a question such as "why is it useful to study complex numbers?" Consider also the interesting question proposed in Harel's chapter—why do mathematicians want to introduce a solution to the equation $x^2 = 1$ but not $x + 1 = x$? Would a successful answer to these questions involve a reference to the proof that was produced? Or would the efficacy of the instruction be determined through students' collective activity as they worked through the instructional sequence? If so, what types of classroom practice could serve as evidence that the learning goal was achieved?

Prioritizing Pedagogical Goals

In each of the chapters, the authors identify pedagogical goals that can be formed by studying mathematicians' practice. We posit that for the mathematics education community to develop better methodological standards for MPP work, we must recognize the implicit arguments by which such knowledge goals are identified and endorsed. We argue that the learning goals in these chapters rely on some form of the following implicit warrant: *these meta-conceptions of proving practice have helped the researcher (and possibly readers) think about mathematical and pedagogical issues, so teachers might benefit from such knowledge as well*. This (clearly tentative) warrant provides plausible evidence that the learning goals could be valuable, though gaining evidence about the effects of such knowledge on future teaching practice is incredibly challenging.

What the implicit warrant does not address is the relative value of the identified meta-conceptions compared with the myriad of other learning goals that could be adopted in mathematics classrooms and pre-service teacher training. Hanna attends to whether students can see proofs that unify bodies of mathematical theory. Harel attends to the quality of teachers' epistemological justifications for complex numbers. Durand-Guerrier and Tanguay attend to different systematizations of the real number line. Boero et al. attend to the standards of mathematical proof in various centuries. On the one hand, we agree that these learning goals are valuable. All things being equal, we would be pleased if these pedagogical goals were attained. On the other hand, these goals strike us personally as somewhat esoteric and we wonder whether a large portion of our own students would achieve them. This does not deny the contribution of these chapters, since each also contains more modest learning goals: perceiving particular proofs as explanatory, learning the properties of the real line, or recognizing generic examples. Rather we identify these different level learning goals to frame the question: what are the roles of the highest level learning goals for the instructional agenda described or implied in MPP work?

Do we expect a large portion of students to attain those goals? If so, then we must attend more carefully to how valuable that knowledge is compared to the instructional opportunity cost. If we do not expect most students to achieve these goals, then how can we properly qualify their role in the presentation of our learning activities? We anticipate that in these chapters (as in our own work on meta-conceptions), these goals provide a conceptual orientation for teaching mathematical content and learning opportunities for our strongest students. We cannot answer these questions for the authors of these chapters, but we think making the question explicit could advance our field's pursuit of MPP research.

Summary

The chapters that we surveyed each make valuable contributions to the emerging field of MPP research. To advance the contributions of these chapters and future work in this vein, we raised several questions about how such research is presented and assessed. As a field, we generally lack clear methodological criteria for transposing ideas from the history or philosophy of mathematics into the mathematics education literature. Furthermore, we recognize that there are implicit warrants operative in such transpositions that should be identified and addressed in order to provide a rich and productive methodology for future TPP research. We hope the issues that we have raised can lead to a broader discussion on the role of MPP research in mathematics education and the standards by which it should be evaluated.

References

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

Avigad, J. (2006). Mathematical method and proof. *Synthese, 153*(1), 105–159.

Balacheff, N. (2008). The role of the researcher's epistemology in mathematics education: An essay on the case of proof. *ZDM Mathematics Education, 40*(3), 501–512.

Boero, P. (2007). *Theorems in school: From history, epistemology and cognition to classroom practice*. Rotterdam: Sense Publishers.

Dawkins, P. C. (2015). Explication as a lens for the formalization of mathematical theory through guided reinvention. *The Journal of Mathematical Behavior, 37,* 63–82.

Dawkins, P. C. (2017). Helping students develop conscious understanding of axiomatizing. In B. Gold, C. E. Behrens, & R. A. Simons (Eds.), *Using the philosophy of mathematics in teaching undergraduate mathematics* (pp. 133–145). Washington, D.C.: Mathematical Association of America.

Dawkins, P. C. (in press). Student interpretations of axioms in planar geometry. *Investigations in Mathematics Learning,* 1–26.

Dawkins, P. C., & Weber, K. (2017). Values and norms of proof for mathematicians and students. *Educational Studies in Mathematics, 95*(2), 123–142.

De Villiers, M. (1990). The role and function of proof in mathematics. *Pythagoras, 24*(1), 17–24.

Durand-Guerrier, V. (2016). Conceptualization of the continuum, an educational challenge for undergraduate students. *International Journal of Research in Undergraduate Mathematics Education, 2*(3), 338–361.

Feferman, S. (1998). *In the light of logic*. New York, NY: Oxford University Press.

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

Harel, G. (2007). Students' proof schemes revisited. *Theorems in school: From history, epistemology and cognition to classroom practice* (pp. 65–78). Sense Publishers: Rotterdam.

Hersh, R. (1993). Proving is convincing and explaining. *Educational Studies in Mathematics, 24* (4), 389–399.

Horsten, L. (2016). Philosophy of mathematics. *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition). Accessed at: https://plato.stanford.edu/archives/win2016/entries/philosophy-mathematics/.

Jaffe, A., & Quinn, F. (1994). Response to "Theoretical mathematics: Toward a cultural synthesis of mathematics and theoretical physics", by A. Jaffe and F. Quinn. *Bulletin of the American Mathematical Society, 30,* 208–211.

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

Lakatos, I. (1976). *Proofs and refutations*. Cambridge: Cambridge University Press.

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

Mejia-Ramos, J. P., & Inglis, M. (2017). 'Explanatory' talk in mathematics research papers. In *Proceedings of the 20th Conference for Research in Undergraduate Mathematics Education.* Downloaded from: http://sigmaa.maa.org/rume/crume2017/Abstracts_Files/Papers/171.pdf.
Resnik, M. D., & Kushner, D. (1987). Explanation, independence and realism in mathematics. *The British Journal for the Philosophy of Science, 38*(2), 141–158.
Selden, A., & Selden, J. (2003). Validation of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education, 34*(1), 4–36.
Thurston, W. P. (1994). On proof and progress in mathematics. *Bulletin of the American Mathematical Society, 30*(2), 161–177.
Weber, K. (2010a). Proofs that develop insight. *For the Learning of Mathematics, 30*(1), 32–36.
Weber, K. (2010b). Mathematics majors' perceptions of conviction, validity, and proof. *Mathematical Thinking and Learning, 12*(4), 306–336.
Weber, K., Inglis, M., & Mejia-Ramos, J. P. (2014). How mathematicians obtain conviction: Implications for mathematics instruction and research on epistemic cognition. *Educational Psychologist, 49*(1), 36–58.
Zelcer, M. (2013). Against mathematical explanation. *Journal for General Philosophy of Science, 44*(1), 173–192.

Part II
Classroom-Based Issues Related to Proof and Proving

Chapter 6
Constructing and Validating the Solution to a Mathematical Problem: The Teacher's Prompt

Maria Alessandra Mariotti and Manuel Goizueta

Abstract Drawing on the hypothesis that an epistemology of school mathematics is interactively constituted in the classroom, we assume that different epistemological stances may lead students to get differently involved in the production and evaluation of arguments as part of their mathematical activity. Based on a case study, in this chapter we focus on how students exploit teacher's interventions to produce arguments to validate different solutions to a mathematical problem within a problem-solving situation. We show that it may happen that teacher's interventions do not have the intended effect, in spite of their potential to foster students' reflection upon the adequacy of these solutions to the proposed empirical situation. Instead, a particular interpretation of the situation emerges through reflection on the solution ultimately validated by the teacher. We depart from this observation to discuss some aspects of the mathematical culture of the classroom.

Keywords Classroom mathematical culture · Validity · Validity construction Argumentation · Mathematical problem · Teacher intervention

Introduction

A growing number of research studies highlight the role of argumentation in mathematical thinking and, consequently, in teaching and learning mathematics at all school levels. These studies support the idea that mathematical activity does not draw solely on deductive arguments, and that non-deductive arguments play a relevant role in advancing mathematics (Inglis et al. 2007). This has led some researchers to claim that a culture of argumentation is to be developed in the

---

M. A. Mariotti (✉)
Università degli Studi di Siena, Siena, Italy
e-mail: mariotti21@unisi.it

M. Goizueta
Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile
e-mail: mgoizueta@gmail.com

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_6

mathematics classroom, and that it should include practices, knowledge and meta-knowledge in relation to the production and evaluation of arguments as part of the classroom mathematical culture, along with the needed awareness to autonomously control such processes (Boero 2011). In order to foster such a culture, it is necessary to know more about the relationship between the construction of mathematical ideas and argumentation, the difficulties students face when engaging argumentative practices, and the challenges faced by teachers who seek to make argumentation central to the mathematics classroom (Ball et al. 2002). Tackling these issues might help teachers and researchers better understand how a culture of argumentation can be deliberately developed in the classroom. In turn, this might help better understand students' mathematical thinking development and perhaps overcome common prejudices suggesting that young students are not capable of it (see Boero et al. this volume).

Drawing on data coming from an empirical study concerning a probabilistic situation problem, in previous contributions we have argued that an epistemology of school mathematics is interactively constituted in the classroom and that it is necessary to understand how it is shaped by and constraints students' argumentative practices (Goizueta et al. 2014). We have illustrated how different epistemological stances may lead students to get involved differently in the production and evaluation of arguments, and the difficulties students and teacher face when dealing with a variety of arguments in classroom conversations (Goizueta and Mariotti 2015). Here, we focus on a pair of secondary students to account for how different arguments emerge in the process of constructing and validating different solutions to a mathematical problem within a problem-solving situation. We particularly focus on how the teacher's interventions are exploited by the students to advance their mathematical work, and on their consequences for validating the specific solution to the given problem. Through our analysis, we aim at showing that the solution finally proposed by the group does not emerge from the students' understanding of and reflection on the empirical situation. Instead, reflection on the empirical situation occurs *ex post facto* (i.e., after the teacher's final validation of the solution), somehow reversing the expected solving process. In our example, specific features of the empirical situation, previously neglected, are then inferred from the solution ultimately validated by the teacher.

Theoretical Perspective

Following Yackel and Cobb (1996) we consider that doing and learning mathematics relates to participating in particular mathematical cultures. We conceive of the mathematics classroom as a particular social context, in which a particular mathematical culture is developed within and by the interaction between teacher and students. This development is influenced by views of mathematics and of mathematics teaching and learning historically produced within broader cultures, of which students and teacher are also members. Thus, the mathematical culture of the classroom is not a pre-given one, but emerges in relation to socio-historical conditions and shows specific features in relation to its context of development.

Participating in the development of a mathematical culture means producing some mathematical propositional knowledge (e.g., definitions and theorems), but also understanding why and how to justify this knowledge and when and how to use it. This implies developing an epistemological control on the mathematical knowledge at stake. In accordance with Steinbring (2005), we assume that a situated epistemology of school mathematics is constituted in classroom interaction. Moreover, as part of this socially constituted school mathematics epistemology, practices aimed at deciding what to believe and on what basis (i.e., epistemic practices) are interactively developed together with a criterion of mathematical validity. A central consequence of these assumptions is the necessity for interpretative research to reconstruct the situated conditions in which (and from which) the validity of mathematical productions is interactively justified in the classroom.

In this perspective, we see validity as related to contingent constraints that are considered good reasons to support/reject a claim in a particular context of justification, and not necessarily restricted to absolute standards (Goizueta et al. 2014). This implies that validity is not a property of claims themselves, but emerges from the mode they are contextually dealt with. It should be stressed that whatever is considered as a good reason, must not necessarily be explicit or even stateable, nor the individuals must be aware of it in a conscious way, it might be enacted implicitly in successful social participation (Ernest 1998).

In classroom interaction, participants are aware of the asymmetric roles played by the teacher and the students. According to these roles, the teacher represents the discipline of mathematics (Voigt 1995). Within this asymmetric interaction, the interplay between teacher's interventions (requests and evaluations) and students' reactions, mediated by (normally implicit) criteria for selection and acceptance of mathematical productions, fosters the emergence of situated epistemic practices. These practices are expected to evolve in time, as teacher and students engage in constructing and validating mathematical knowledge in increasingly sophisticated ways. Therefore, we conceive of the epistemology of school mathematics as a dynamic system that enables and constrains mathematical activities, which, reflexively, contribute to its constant development.

Like Ernest (1998), we consider that knowledge construction is a social, dialogical human activity. We see social interaction as a means to produce, revise, correct and validate knowledge, and argumentation as the common core of such epistemic practices. It is by offering arguments to support or criticize one's own and other's claims that the dialogical dimension of knowledge construction is deployed. Within this frame, we speak of argumentation to refer to certain aspects of social interaction, and not to some meta-communicative activity that is undertaken *ex professo* to secure the validity of ordinary actions (Krummheuer 1995). An exchange is argumentative whenever it conveys reasons to establish the epistemic status of some claim, either explicitly or implicitly. For instance, while trying to find a solution to a problem, students may perform some calculations that already express reasons to support the validity of a solution. In that case, the process of calculating itself presents an argumentative nature. As Krummheuer, we consider that an exchange can be reconstructed by the participants or by an observer in order to express its argumentative nature in what we call an argument (for a particular, theory-driven way to reconstruct arguments see Pedemonte, this volume). This implies that being argumentative is not a property of texts (written or spoken) themselves, but emerges from the mode they are contextually interpreted.

Participants, Task and Data Collection

The experiment we report about was part of the larger Ph.D. study of the second author (Goizueta 2015), in which two different experimental situations, regarding two different teachers and classrooms, were considered. The participants in the experimental situation we report here were twenty-one 14/15-year-old students and their teacher in a secondary school mathematics classroom in Catalonia, Spain. The data we present in this report comes from two lessons. It was a problem-solving setting, with time for small group work and whole-class discussion, what constituted a common working situation for the class. The researchers suggested the following task:

> Two players are flipping a coin in such a way that the first one wins a point with every head and the other wins a point with every tail. Each is betting €3 and they agree that the first to reach 8 points gets the €6. Unexpectedly, they are asked to interrupt the game when one of them has 7 points and the other 5. How should they split the bet? Justify your answer.

The task was intended as an introduction to probability and aimed at setting grounds for discussing basic notions such as random game, (non) equiprobable events, etc. This was the first contact for the group with probability in the context of school mathematics. We were aware of the high complexity of the task, related to the lack of shared school probabilistic notions (Wilensky 1997), and expected the students to tackle the problem drawing upon their arithmetical knowledge and tools. The choice of the content topic was inspired by Ruthven and Hofmann (2013), who consider probability as especially appropriate for making complementary epistemic approaches emerge in secondary mathematics.

In order to solve the problem, it is necessary a "translation" effort between the empirical situation and mathematics, in both directions. That is, on the one hand, some empirical elements and relationships that define the situation must be interpreted mathematically; and on the other hand, the mathematical results obtained must be interpreted in empirical terms to give a meaningful answer to the problem. The validity of a solution to the problem is related to how well it represents the empirical situation (as it is interpreted by the solver), thus in order to secure such validity, it might be necessary to provide an argument explicitly referring to the relationship between the solution and the situation. It is precisely through the assessment of the representativeness of different solutions that different interpretations of the empirical situation might emerge and, consequently, a variety of solving processes can be scrutinized. The comparison of competing solutions by the students is expected to foster different arguments to support or reject them. It is within such activity that basic notions of probability are expected to emerge and be discussed in order to account for different interpretations of the situation and in relation to different solutions.

The considerations above, together with previous results concerning the use of this problem with secondary students (Paola 1998) and studies related to different historical solutions given by mathematicians (Fenaroli et al. 2014), suggested this problem might be a good candidate to promote a rich argumentative environment in the classroom and to foster the discussion of mathematical, but also meta-mathematical issues. The novelty of the task was expected to prevent students from using mechanical approaches based on well-established solving strategies. In line with the pedagogical approach by Zaslavsky (2005) and Brown (2014), we expected uncertainty about the validity of competing solutions to generate a reflexive environment in the classroom, as well as the need to produce a variety of arguments.

For data collection, three small groups were videotaped and written protocols were collected. The obtained data were coded with the aid of qualitative data analysis software. We conducted interpretative analyses following a constant comparison approach on the codified data (Strauss and Corbin 1998). In a later stage, based on the triangulation with two other researchers familiar with the theoretical perspective (i.e., seeking consensus in relation to the interpretation of data), these analyses allowed us to construct a number of themes in order to account for the mathematical activity in the classroom. These themes are narrative texts that articulate the most relevant aspects of the analyzed data in accordance with the theoretical perspective adopted (van Manen 1990). Through this process, the relationship between specific teacher's interventions and the construction of the validity of students' mathematical work emerged as a relevant aspect of the mathematical activity in the classroom.

The Teacher's Interventions and the Students' Mathematical Work

During the two lessons, Dan, the teacher, often emphasizes the need to justify any proposed solution to the problem. He's initial comment, when introducing the problem, exemplifies this.

001 Dan: You must give an argument. You cannot say: this. No, [you must say] this because of some reason, some argument. (…) You must justify your answer.

This and similar comments along the lessons could be thought of as attempts to establish justification as an essential feature of any acceptable solution to the problem. However, Dan does not explicitly elaborate on what counts as justification in this context. While interacting with the groups, Dan tends, on the one hand, to make explicit evaluations of the students' solution's validity; on the other hand, he offers hints that relate the solution he expects to empirical aspects of the situation proposed in the problem's wording. Although his interventions could foster critical reflections about the students' own solutions in relation to how well they represent the empirical situation, we do not observe among the students explicit reflections in this direction. Instead, students exploit Dan's interventions only to discard specific solutions that were negatively evaluated, identify specific expected features, and construct new solutions accordingly. Following this hint-guidance by the teacher, the three observed groups finally arrive to the expected (probabilistic) solution, whose validity is ultimately confirmed by the teacher. It is just after the teacher's validation that we can observe within the groups explicit reflections on the adequacy of the solution to the empirical situation and its dynamic. That is, once the teacher accepts a particular solution, the empirical situation is interpreted and conceptualized accordingly.

To illustrate this process, which was observed in all three groups, in the following we focus on the group composed by Tess and Jay and analyze the transcripts of students' interactions. Within the transcripts, numbers on the first column indicate speaking turns. Group work on the first and second lesson, and whole group discussion where numbered independently.

Exploiting the Teacher's Hints to Progress in the Mathematical Work

In early stages of the approach to the problem, the students explore a first mathematical solution, which we could paraphrase: 'if by winning 8 points a player gets €6, for each point won a player should get €0.75'. Jay falsifies this solution offering a counterexample that draws on a variation of the problem (cf. Komatsu et al. this volume, who elaborates on the relevance of counterexamples in advancing proof construction and validation).

030	Tess:	What if we do: if eight winning tosses are worth six euro, seven winning tosses… how much are they worth?
031	Jay:	How did you do that? What did you do?
032	Tess:	Proportion [uses calculator]. This one gets five point twenty five euro.
033	Jay:	But then imagine that the other one won seven as well. 'Cause he could. Then he should also get the same and no, no…
034	Tess:	Right, that's true! This is wrong.

Jay's argument [33] and Tess' agreement [34] indicate that they consider that numerical variations of the problem might inform about the validity of the solution, what suggests that they expect a valid solution to be applicable in general. Jay exploits a particular variation to notice that the solution does not adequately represent an empirical datum, namely, that six euros must be distributed. The solution is thus falsified by noticing that it does not adequately represent the empirical situation. The group then comes up with a new solution, corresponding to distributing the money proportionally to the points won, and ascertains that the amounts of money distributed add up to six in two different numerical variations, namely when the game has stopped and the score is four points to three, and when the game stops tied at seven points. The students take this fact as support for the new solution's validity. After this inductive argument, the group presents the numerical results to the teacher as the solution to the problem; what indicates a high level of confidence in the solution's validity. When reporting their work to Dan, the students only propose the numerical results (€3.5 for the winning player and €2.5 for the opponent) without explaining the process carried out to obtain and validate them. Both the argument for the falsification of the first solution (a counter example implicitly drawing on the solution's general applicability) and the inductive argument supporting the second one are omitted in the interaction with the teacher.

Dan's reaction does not facilitate the emergence and discussion of the students' arguments. He rejects the solution presented by the group and offers new hints related to specific features of the expected solution.

068	Dan:	If I was the winning player I would not be satisfied. Would you? I wouldn't.

Later on, when Tess insists on the proportional model:

146	Tess:	But why is it not right, Dan?
147	Dan:	Because it is not. It's not about what you have done, but what has yet to be done. The possibilities that are yet to be done. (…) You have to see the future. What can happen from seven? Who could win? What are the possibilities for one player and the other? Do some schemes, do some graphs. Check it out. [leaves]

Dan interventions explicitly convey his negative evaluation of the numerical results' (and hence implicitly of the solution's) validity [68, 147]. He hints the students about specific features of the expected solution: the winning player should get more money [68] and what can still happen and not to what has already happened during the game must be considered [147]. It is likely that Dan's aim is to offer some ideas to reflect on the empirical situation, its relevant elements and their relationships, in order to construct the expected solution. His initial (and continuous) emphasis on justification reinforces this interpretation. Dan's hints could lead students to reflect on the relationship between the proposed solution and the empirical situation by addressing crucial questions, e.g., why should the winning player get more money? or why what can still happen should be considered and not what has already happened? In turn, such reflections could potentially lead the students to new understandings about the empirical situation and, through them, to show that the proportional solution does not adequately represent it. On such basis, more representative solutions could abductively emerge. Nevertheless, in none of the groups we observe such reflections. Instead, students operationalize Dan's hints to identify new sets of numerical data, with which to calculate new numerical results in line with perceived teacher's expectations. The following conversation between Tess and Jay illustrates this way of exploiting the teacher's hints.

178	Tess:	See? There are four options. Three give the victory to this and one to this.
179	Jay:	(…) Yes, yes, yes, of four possibilities, player one has three.
180	Tess:	Of four possible final scores, three are player one's victories and just one [is] player two's [victory]. That's it!

Here, following Dan's hints, Tess considers what can happen in terms of possible final scores and uses them to propose a new distribution: each player receives an amount proportional to the number of favorable final scores [178, 180]. After doing the calculations, she notices that this solution gives more money to the winning player than the previous one (€4.5). The construction of this new solution to the problem and its purported validity seem to be mainly founded on the perceived accordance to Dan's hints rather than sprouting from students reasoning, as was for the previous cases. This interpretation is confirmed the next lesson, when the group reports to the teacher the new numerical results.

015	Tess:	Dan, would this be OK? [notes: "6 × ¾ = 4.5€" and "6 × ¼ = 1.5€"].
016	Dan:	I think I wouldn't be satisfied with that. If I was this [winning player]. (…)
019	Jay:	But, we focused on the remaining tosses, and not on the past ones. (…)
022	Dan:	Let's see, not all possibilities have the same probability. Do you understand? Here you have three and one, but they are not equally possible. Do you understand? (…) it is not three and one. Well, it is three and one, but this one… The three and the one might not have the same weight. Do you understand?

As before, the students do not present the rationale behind the emergence of the new solution, nor does Dan propitiate its discussion. Jay briefly refers to it [19], but seemingly to stress that they followed Dan's indications, and not to justify the underlying reasoning. The teacher negatively evaluates the numerical results [16] and provides a new hint that could foster students' reflection on empirical aspects of the situation [22] (e.g., the likelihood of different final scores). Later on, Dan will suggest the group to construct a tree representation of the possible evolution of the game, a hint that will allow Tess to produce a new solution, corresponding to the expected probabilistic one.

Exploiting the Teacher's Validation to Interpret the Empirical Situation

057	Tess:	Player two has twelve point five percent of possibilities of winning.
058	Jay:	How did you get that?
059	Tess:	[tree representation on her notes (Fig. 6.1)] Fifty percent of fifty percent is twenty-five percent. Fifty percent of twenty-five percent is twelve point five percent. So this one has fifty percent plus twenty-five percent plus twelve point five percent. He has eighty-five point five percent. (…) So we should multiply six by eighty-seven point five percent. (…)
062	Dan:	I like that better. [leaves]

Tess knows that getting heads or tails is equally likely; this seems clear from what she says [57, 59]. She uses this idea, expressed as a percentage, together with the tree representation suggested by the teacher (Fig. 6.1) to describe how "possibilities" evolve with each toss [59]. By adding up these possibilities she obtains two percentages (corresponding to each player's probability of winning) that uses to distribute the money: 87.5% of the bet for the winning player and 12.5% for the opponent, what correspond to €5.25 and €0.75€. This new solution and numerical results are in line with the teacher's hints: the winning player gets more money and each final score has a different "weight" (expressed as a percentage). From what Tess says and does, it is not possible to understand in what sense is she considering the relationship between this solution and the empirical situation. She might be mainly manipulating the numerical data to obtain a solution in line with the teacher's hints. In any case, the underlying argument justifying the validity of the

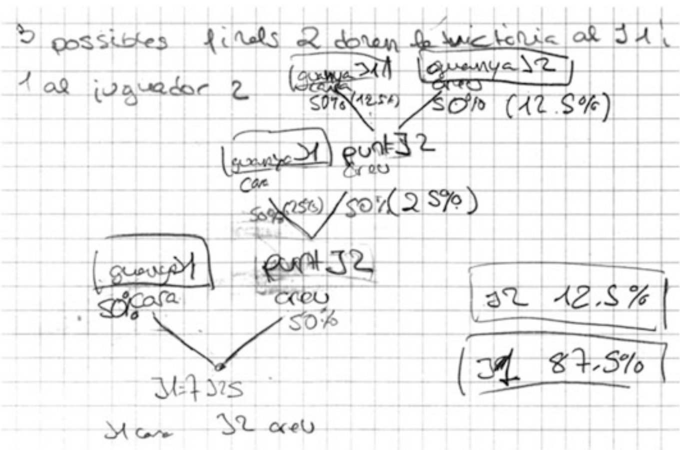


Fig. 6.1 Tree representation of the possible game evolution

solution to the problem seems to draw on its agreement and on the agreement of the numerical results with the teacher's expectations. This time Dan expresses satisfaction with the student's solution [62].

It is just after what is taken as the teacher's validation of the solution and numerical results that the students autonomously reflect on their relation to the empirical situation.

070	Jay:	So, clearly, they could have tossed the coin any number of times before.
071	Tess:	Yes, previous tosses do not matter. It is about the remaining tosses. But then, the number you must get to doesn't matter either.
072	Jay:	What?
073	Tess:	For example, if the goal is ninety-nine points and one has won ninety-eight and the other one ninety-five… It doesn't matter.
074	Jay:	So, what we did before was not right.
075	Tess:	No. That was poop.
076	Jay:	I don't understand why that can't be as well. Just because Dan says that, but…

Reflecting on the solution, Jay seems to realize that the points won are irrelevant to calculate the amounts of money corresponding to each player [70]. Tess complements Jay's inference by noticing the irrelevance of the total of points needed to win the game [71, 73]. This is a key feature of the probabilistic approach that, as we observe, does not emerge from the empirically grounded reflection on and understanding of the situation. Instead, this feature is inferred from the solution that has been validated by the teacher. Reflection on and understanding of the empirical situation does not occur as a means to produce the 'correct answer', but follows the questioning of the teacher's validation of the solution (076 "Just because Dan says that, but …"). Thus, validity does not emerge from considering the solution's representativeness of the empirical situation. Instead, once validated by the mathematical authority of teacher, the students seem to look for an interpretation of the situation that could accommodate features inferred from the warranted solution. That is, the empirical situation is interpreted and conceptualized *ex post facto*, in accordance with the solution validated by the teacher.

Jay's last comments [74, 76] and the fact that Tess does not address them suggest that, despite having arrived to the expected (probabilistic) solution, the probabilistic nature of the empirical situation is not yet clear to the students. It seems that, for these students and this specific task, the teacher's authoritative validation of the solution does not suffice to achieve personal conviction. This might indicate that, although the teacher's mathematical authority suffices to sanction the validity of the solution, his interventions lack the explanatory power (cf. Hanna, this volume) needed in this case to help the student relate the solution to specific features of the empirical situation. Moreover, the students do not discuss this issue with the teacher, nor do they further reflect on the different solutions and their relation to the empirical situation to clarify it.

During whole group discussion, and after the probabilistic solution was presented by a student and agreed by the whole group, Dan elaborates on how some students proceeded for its construction.

023 Dan: Look, some of you have reversed the reasoning. You have to do some reasoning in order to arrive to the solution. But not from the solution justify the argument. Is that clear?

This interventions by Dan suggests that he has observed that some of the groups reached a probability-related interpretation of the situation just after the solution "was arrived at". However, since none of the students replied to Dan's final question, nor he further discusses students' approaches to solve the problem, it is not possible to know what is Dan referring to or how the students interpret what is said. Nevertheless, it seems clear that Dan considers the approach of some groups to the construction of the probabilistic solution as problematic. In view of our analysis of Jay and Tess's approach, Dan's intervention is relevant, but it is seemingly not sufficient to illuminate crucial aspects of mathematical activity (e.g., the use of abductive reasoning to infer a mathematical solution that represents the empirical situation). Although this looks like a good occasion for Jay to raise his concern about the validity of the solution preferred by the teacher, he does not intervene. It is possible that he has reached a new understanding about the empirical situation after his classmate's presentation, allowing him to convincingly justify the probabilistic solution's validity. But it might also be the case that he does not perceive raising his concerns about the justification of the agreed solution as a mathematically relevant part of the ongoing mathematical activity.

Final Discussion

Our analysis shows that the teacher's interventions and didactical choices are critical to advancing and validating mathematical production, but also that this might happen in unintended ways. This resonates with the results of Pedemonte (this volume), who points out particular aspects of teacher's arguments that might be critical in helping students develop their own arguments to tackle proof tasks.

Although Dan's interventions could foster empirically grounded reflections about the representativeness of the proposed solutions, and are seemingly intended that way, we do not observe among the students explicit reflections to this regard. The students exploit the teacher's interventions to discard rejected solutions, and operationalize the offered hints to construct new ones satisfying perceived expectations. In this process, the crucial mathematical activity consisting of systematically reflecting on the relationship between the solution and the empirical situation seems to be neglected or at least not sufficiently exploited by the students. The mathematical authority of the teacher emerges as a major support for the argumentative construction, rejection and validation of successive solutions. As a consequence, features of the empirical situation relevant to its probabilistic interpretation (e.g., probabilities' independence of past events) do not reflectively emerge as criteria to warrant validity. Instead, these features are (deductively) inferred from the solutions preferred by the teacher, so that a particular interpretation of the empirical situation comes out as a (deductive) consequence of the solution validated by the teacher.

We do not claim that such way of proceeding is mathematically meaningless; in fact, considering the deductive consequences of mathematical solutions is a relevant practice within experimental sciences in order to test solutions' validity. But we argue that developing an understanding of the epistemological implications of this and other epistemic practices is crucial for the construction of a culture of argumentation as part of students' mathematical and scientific literacy. In the example discussed above, teacher's interventions and the way they are exploited do not seem to help the students address the relationship between validity and representativeness. This way of exploiting the teacher's interventions might be the students' way to adjust to what they perceive as expected from them as mathematics learners in that particular mathematics classroom culture, namely to produce the solution expected by the teacher; a common feature of many mathematics classrooms. If this is the case, important opportunities for constructing mathematical and meta-mathematical knowledge are missed due to representations of what is (in) adequate in the mathematical culture of the classroom. If we expect argumentation to be central to students' mathematical work and to constitute a means for the construction of mathematical and meta-mathematical knowledge and the development of mathematical thinking, justification must be a relevant goal, a valued activity and an object of explicit reflection within the classroom mathematical culture.

To help teachers develop such mathematical culture in the classroom, Boero et al. (this volume) suggest that it is worth planning and implementing pre-service teacher courses centered in mathematical argumentation. Argumentation oriented courses might help future teachers design and plan activities in which students take responsibility for the validity of mathematical productions and its discussion, and in which the explanatory power of mathematical arguments is taken into consideration (Hanna, this volume). Such activities might help the deliberate construction of an epistemology of school mathematics in line with pedagogical objectives. To this regard, long term studies aimed at understanding how an epistemology of school mathematics is interactively developed as part of the classroom mathematical culture are essential, both to understand how to guide this development and to inform future teachers' education accordingly.

Acknowledgements The second author's participation in this study has been partly funded by the Spanish Grant EDU2012-31464 and the Catalan Grant SGR2014-972.

References

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

Boero, P. (2011). Argumentation and proof: Discussing a "successful" classroom discussion. In M. Pytlak, T. Rowland, & E. Swoboda (Eds.), *Proceedings of CERME7* (pp. 120–130). Rszéskow, Polonia: ERME.

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

Ernest, P. (1998). *Social constructivism as a philosophy of mathematics*. Albany, NY: State University of New York Press.

Fenaroli, G., Guala, E., Goizueta, M., Paola, D., & Sanna, G. (2014). Il problema delle parti per una introduzione al pensiero probabilistico. *L'Insegnamento della Matematica e delle Scienze Integrate, 37*(6 A + B), 573–584.

Goizueta, M. (2015). *Epistemological aspects of argumentation in the mathematics classroom* (Doctoral thesis). Autonomous University of Barcelona, Barcelona, Spain.

Goizueta, M., & Mariotti, M. A. (2015). Constructing validity in classroom conversations. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the 9th Congress of the European Society for Research in Mathematics Education* (pp. 128–134). Prague, Czech Republic: Charles University in Prague.

Goizueta, M., Mariotti, M. A., & Planas, N. (2014). Validating in the mathematics classroom. In S. Oesterle, P. Liljedahl, C. Nicol, & D. Allan (Eds.), *Proceedings of the Joint Meeting of PME 38 and PME-NA 36* (Vol. 3, pp. 169–176). Vancouver, Canada: PME.

Inglis, M., Mejía-Ramos, J. P., & Simpson, A. (2007). Modeling mathematical argumentation: The importance of qualification. *Educational Studies in Mathematics, 66*, 3–21.

Krummheuer, G. (1995). The ethnography of argumentation. In P. Cobb & H. Bauersfeld (Eds.), *The emergence of mathematical meaning: Interaction in classroom cultures* (pp. 229–269). Hillsdale, NJ: Lawrence Erlbaum Associates.

Paola, D. (1998). Il problema delle parti. Prassi didattica e storia della matematica. *La didattica delle scienze, 198,* 31–16.

Ruthven, K., & Hofmann, R. (2013). Chance by design: Devising an introductory probability module for implementation at scale in English early-secondary education. *ZDM-Mathematics Education, 45,* 409–423.

Steinbring, H. (2005). *The construction of new mathematical knowledge in classroom interaction: An epistemological perspective*. New York: Springer.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing Grounded Theory*. Thousand Oaks, CA: Sage.

van Manen, M. (1990). *Researching lived experience: Human science for an action sensitive pedagogy*. London, UK: Althouse.

Voigt, J. (1995). Thematic patterns of interaction and sociomathematical norms. In P. Cobb & H. Bauersfeld (Eds.), *The emergence of mathematical meaning: Interaction in classroom cultures* (pp. 229–269). Hillsdale, NJ: Lawrence Erlbaum Associates.

Wilensky, U. (1997). What is normal anyway? Therapy for epistemological anxiety. *Educational Studies in Mathematics, 33,* 171–202.

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

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

Chapter 7
Addressing Key and Persistent Problems of Students' Learning: The Case of Proof

Andreas J. Stylianides and Gabriel J. Stylianides

Abstract Research has provided a strong empirical and theoretical basis about major difficulties students face with proof, but it has paid less attention to the design of interventions to address these difficulties. In this chapter we highlight the need for more research on classroom-based interventions in the area of proof, and we discuss what might be important characteristics of interventions that specifically aim to address key and persistent problems of students' learning in this area. In particular, we make a case for interventions with the following three characteristics: (1) they include an explanatory theoretical framework about how they "work" or "can work" in relation to their impact on students' learning; (2) they have a narrow and well-defined scope, which makes it possible for them to have a relatively short duration; and (3) they include an appropriate mechanism to trigger and support conceptual change. Although our discussion of these characteristics focuses on the area of proof, the characteristics can be applicable also to interventions that aim to address key and persistent problems of students' learning in other areas.

Keywords Cognitive conflict · Proof · Intervention · Misconception
Learning · Design-based research

Introduction

Several scholars have expressed concern that mathematics education research has played an inadequate role in supporting improvement of classroom practice, especially improvement of students' learning of mathematics (e.g., Ruthven and Goodchild 2008; Stylianides and Stylianides 2013; Wiliam and Lester 2008). Students' learning of proof is no exception to this trend (Stylianides et al. 2016, 2017).

A. J. Stylianides (✉)
University of Cambridge, Cambridge, UK
e-mail: as899@cam.ac.uk

G. J. Stylianides
University of Oxford, Oxford, UK

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_7

According to a recent review of the state of research in the area of proof, Stylianides et al. (2017) concluded that, while research has provided a strong empirical and theoretical basis about different difficulties students face with proof, it has paid less attention to the design of interventions that will aim to address these difficulties. This state of affairs undermines efforts to elevate the status of proof in students' mathematical experiences, because some significant problems of students' learning in the area of proof remain without solutions and create obstacles to further learning in this area. Indeed, it would be unrealistic to expect that, if left to their own devices, individual teachers, textbook authors, or other stakeholders would be able to successfully navigate a pedagogically arduous territory and design appropriate and adequate learning experiences to help students overcome significant difficulties they face in the area of proof.

Thus, there is a pressing need for researchers to embark on the design of classroom-based interventions in the area of proof, *specifically interventions that would aim to address key and persistent problems of students' learning*. One form these problems can take, which is the form we focus on in this chapter, is that of common and hard-to-remediate student misconceptions about key aspects of proof. An example of such a misconception is that a single counterexample is not sufficient to refute a false mathematical generalization (e.g., Mason and Klymchuk 2009; Zaslavsky and Ron 1998). Another misconception, to which we return later in the chapter for illustration purposes, is that a few confirming cases are enough to establish the truth of a mathematical generalization (e.g., Morris 2007; for a review, see Harel and Sowder 2007). The latter misconception relates to one of the "solid findings" of research on mathematics teaching and learning identified by the Education Committee of the European Mathematical Society (2011): that "*many students provide examples when asked to prove a universal statement*" (p. 50, emphasis in original), i.e., students tend to consider that empirical arguments are proofs of mathematical generalizations. The Education Committee noted the following about this misconception that help explain why we consider it to be an example of a key and persistent problem of students' learning in the area of proof:

> [C]onsiderable evidence exists that many students rely on validation by means of one or several examples to support general statements, that this phenomenon is persistent in the sense that many students continue to do so even after explicit instruction about the nature of mathematical proof, and that the phenomenon is international … (Education Committee of the European Mathematical Society 2011, pp. 50–51)

Recognizing the need for more research on classroom-based interventions in the area of proof raises the following question: *What might be some important characteristics of interventions that aim to address key and persistent problems of students' learning in the area of proof?* In this chapter we discuss this question, drawing on our own intervention-based research in the area of proof and other relevant literature. Our primary aim is to draw attention to this significant but inadequately researched topic, and to discuss ideas that can serve as a starting point for a conversation in the field about how to organize classroom-based interventions in the area of proof so as to gain leverage in addressing more programmatically the serious problems that students face with proof.

Before we review, as a background for the chapter, selected prior research on classroom-based interventions in the area of proof, we clarify our use of key terms, drawing on Stylianides and Stylianides (2013, p. 334). Although the term "intervention" has been used in different ways, our use of it in this chapter is similar to its standard use in medicine where an intervention denotes "action taken to improve a situation" (Stevenson and Lindberg 2012). In our case, the situation in need for improvement is an aspect of students' learning in the area of proof that deviates from conventional knowledge, while the action taken to improve the situation is purposefully designed classroom instruction. Also, we use the term "classroom" broadly to denote a formal learning setting at any level of education: elementary, secondary, or university (including teacher education). Finally, we use the term "learning" in a broad way, too, to include not only cognitive but also affective aspects of students' mathematical experience.

Background: Existing Research on Classroom-Based Interventions

Although there are notable examples of research studies on classroom-based interventions in the area of proof at least as early as the 1930s (Fawcett 1938), the number of these studies is small and acutely disproportionate to the number of studies that documented problems of students' learning. In this section we briefly present five recent classroom-based interventions that were all satisfactorily successful in promoting their intended goals in the area of proof. The first two interventions focused on the secondary school level (in the domain of geometry) and the other three on the university level. For a more detailed discussion of these and other related interventions the reader can refer to Stylianides et al. (2017). Also we are currently guest-editing a special issue in *Educational Studies in Mathematics* on the same topic (Stylianides and Stylianides 2017).

Mariotti's (2013) intervention aimed to introduce 15–16-year-old students in Italy to the deductive approach in geometry, a goal that is often difficult for instruction to promote. The intervention extended over a 2-year period and used the dynamic geometry environment Cabri-Géomètre (hereafter, Cabri). The Cabri tools initially available to the students corresponded to the straightedge and compass tools used in the traditional paper-and-pencil environment. As the students developed different geometrical constructions, the Cabri menu was expanded to include new commands, which then became theorems available for use in subsequent constructions. With the teacher's mediating role, the notion of proof in the intervention served as a tool that ensured the validity of new constructions based on the available commands and as a key aspect of the social contract in the classroom whereby constructions had to be justified before they became theorems.

Jahnke and Wambach's (2013) intervention with eighth-grade students in Germany aimed to develop students' understanding that a proof is based upon certain assumptions, another important but hard-to-achieve learning goal. The intervention extended over eight geometry lessons and was contextualized in the attempts of ancient Greeks to model the path of the sun, the so-called "anomaly of the sun." The students were asked to put themselves in the position of the ancient astronomers and to assume that available to them were only the methods and tools that were known at the time. These restrictions were similar to the restrictions imposed on the Cabri tools available to the students in Mariotti's (2013) study and were an important factor to students becoming more conscious of the role of assumptions in building a deductive theory.

In Stylianides and Stylianides (2009) we reported on an intervention that we developed in a 4-year design experiment in an undergraduate mathematics course for preservice elementary teachers in the United States. The intervention lasted less than three hours and aimed to help preservice teachers begin to overcome the misconception we described in the Introduction—namely, that empirical arguments offer secure methods of validating mathematical generalizations—and to see an "intellectual need" (Harel 1998) to learn about secure methods of validation (i.e., proofs). The intervention involved the implementation of a purposefully designed task sequence as well as two deliberately engineered "cognitive conflicts" and associated "conceptual awareness pillars" (both notions are discussed in the following section) that motivated and supported stepwise progressions in preservice teachers' knowledge about proof along a pre-specified learning trajectory. The learning trajectory began from a naïve empirical conception (Balacheff 1988), continued to a crucial experiment conception (ibid.), and ended with a non-empirical conception (Stylianides and Stylianides 2009). An appropriately adapted version of the intervention was subsequently implemented, with similarly promising results, in a high attaining secondary mathematics classroom in England (Stylianides and Stylianides 2014b).

Harel (2001) reported on a two-week intervention that was part of a broader teaching experiment in an elementary number theory course for lower secondary preservice mathematics teachers. The intervention aimed to teach proof by mathematical induction in a way that addressed major deficiencies of traditional teaching of this proof method, drawing on a system of pedagogical principles elaborated in Harel (2010). According to Harel (2001), the most important finding of his intervention was that students changed their "ways of thinking, primarily from mere empirical reasoning – in the form of result pattern generalization – into transformational reasoning – in the form of process pattern generalization" (p. 206).

Hodds et al. (2014) reported on a series of experiments they conducted with undergraduate mathematics students to investigate the effect of an intervention, in the form of a generic self-explanation training, on students' proof comprehension. The training, which took less than 20 min of individual study, aimed to address some common limitations in undergraduate students' proof comprehension strategies by focusing students' attention on logical relationships within a proof. Following the positive findings of two experiments under lab conditions, the research team investigated the effect of the intervention in a genuine pedagogical setting and found a similar positive effect that persisted for at least three weeks.

Important Characteristics of Interventions to Address Key and Persistent Problems of Students' Learning

The review in the previous section shows that the body of research on classroom-based interventions in the area of proof has produced some notable results though, of course, much more needs to be done in this area. The review shows further that, while all interventions aimed to promote important and hard-to-achieve learning goals in the area of proof, only a few of them aimed specifically to address key and persistent problems of students' learning, such as common and hard-to-remediate student misconceptions about key aspects of proof. This observation is not a criticism of the aims of different studies; rather, our intention is to point out the scarcity of research on the particular aim we focus on in this chapter.

In this section we discuss three characteristics that we consider important for researchers to take into account as they design classroom-based interventions to address key and persistent problems of students' learning. The three characteristics are summarized in Table 7.1 with an indication also of whether we consider each characteristic to be desirable or essential given the particular kind of intervention we focus on herein. The reader will notice that the summary of the three characteristics in Table 7.1 does not mention the word "proof." Indeed, we consider that

Table 7.1 Important characteristics of interventions that aim to address key and persistent problems of students' learning (in the area of proof but also more broadly)

Characteristic	Description	Desirable or essential?
1. An explanatory theoretical framework	An account for how the interventions "work" or "can work" and thus an identification of key features of the interventions to which their impact can be attributed	Essential for, otherwise, the interventions' impact can be due to idiosyncratic factors and their successful adaptation for use in other contexts may be unlikely
2. A narrow and well-defined scope	The interventions target few well-defined learning goals, which can allow them to have a relatively short duration	Desirable, though it is hard for interventions that do not have this characteristic to fully meet Characteristic 1
3. An appropriate mechanism to trigger and support conceptual change	The interventions include a mechanism that can help trigger and support conceptual change in relation to the problem of students' learning (e.g., misconception) targeted by the intervention	Essential for, in the absence of such a mechanism, the problem of students' learning will persist

the characteristics can apply also to interventions in areas other than proof. However, given our focus in this chapter on proof, our discussion of the three characteristics will center around this specific mathematical area.

In the rest of this section, we first present each characteristic and then we exemplify it by drawing on our own research in the area of proof. We clarify that do not suggest that interventions (past or future) that do not meet these characteristics are lacking; these interventions are simply of a different nature from the one we focus on in this chapter.

Characteristic 1: An Explanatory Theoretical Framework

We consider essential for research that produces successful interventions in addressing key and persistent problems of students' learning in the area of proof (or in another area of mathematics) to also produce explanatory theoretical frameworks to account for how the interventions achieved their positive outcomes. Designing successful interventions is certainly welcome and a major accomplishment in itself, but the broader contribution of these interventions would nevertheless be limited without an understanding of how the interventions "work" or "can work" to support student learning.

The explanatory theoretical frameworks we call for are essentially "design theories" in a sense that is typical of the ones found in studies using *design experiment methodology* (e.g., Cobb et al. 2003; Design-Based Research Collective 2003). We follow Cobb et al. (2003) in characterizing these theories as "relatively humble in that they target domain-specific learning processes" (p. 9), in our case important but hard-to-teach learning processes in the area of proof, and we view an intervention as the "instructional engineering" (Stylianides and Stylianides 2014b) by which a research team tries to generate these learning processes in the mathematics classroom. Cobb et al. illustrated the meaning of a design theory as follows:

> For example, a number of research groups working in a domain such as geometry or statistics might collectively develop a design theory that is concerned with the students' learning of key disciplinary ideas in that domain. A theory of this type would specify successive patterns in students' reasoning together with the substantiated means by which the emergence of those successive patterns can be supported. (Cobb et al. 2003, p. 9)

The "successive patterns in students' reasoning" mentioned in the quotation can correspond, for example, to the various milestones in a learning trajectory that are achieved by students in a classroom during the implementation of an intervention. The "substantiated means" (also mentioned in the quotation) can be, for example, the actions of the teacher while implementing the intervention and the curricular resources (e.g., mathematics tasks) used in the intervention, as well as the interactions between these and other components of instruction that are in play during the intervention (for different frameworks to analyze these interactions in the area of proof and beyond, see Stylianides 2016b).

In the absence of an explanatory theoretical framework one cannot exclude the possibility that the impact of a successful intervention was due to idiosyncratic factors, including a fortunate coincidence of particular teacher and student characteristics with the coalition of favorable classroom circumstances. It is partly for this reason that one should expect a research study showing that an intervention "works" or "can work" to include what Greeno (2006) described as "analytical hypotheses about what features of the design were responsible for the successes and limitations of its accomplishments in the circumstances that were in place in the study, especially ways in which these circumstances were related to design principles" (p. 799). In other words, the research team should offer a rational, evidence-based account of how the substantiated means that were in play during the intervention supported the successive patterns in students' reasoning that emerged from the intervention. Once a design theory about how the intervention plays out, or can play out, in the classroom is firmly in place thereby specifying the theoretically essential components of the intervention, decisions can be made about what aspects of the intervention can be modified and how, and what others should stay invariant, so as to minimize the risk of poor results when using the intervention in new contexts (Greeno 2006; Yeager and Walton 2011).

Our own intervention-based research was conducted within the frame of design experiment methodology, which, as we alluded to earlier in this chapter and explained elsewhere (Stylianides and Stylianides 2013), is ideally suited to serve the following two goals: designing interventions that "work" or "can work" in promoting student learning in real classroom settings (goal 1), and developing theory to explain the mechanisms that supported student learning thereby creating knowledge that can be useful beyond the local setting wherein it was created (goal 2). None of these goals can be underestimated or overlooked, and indeed, in the context of design-based research, it is difficult to see the two goals in isolation of each other. While "[d]esign experiments are conducted to develop theories [see goal 2], not merely to empirically tune 'what works' [see goal 1]" (Cobb et al. 2003, p. 3), it is also true that "design-based research that advances theory but does not demonstrate the value of the design in creating an impact on learning in the local context of study has not adequately justified the value of the theory" (Barab and Squire 2004, p. 4). The reference to a study's "local context" in the previous quotation may raise the question about whether the study's findings can be characterized as local and thus of little broader interest. This is not the case:

> Although, as a practical matter, a design experiment is conducted in a limited number of settings, it is apparent from the concern for theory that the intent is not merely to investigate the process of supporting new forms of learning in those specific settings. Instead, the research team frames selected aspects of the envisioned learning and of the means of supporting it as paradigm cases of a broader class of phenomena. (Cobb et al. 2003, p. 10)

The intervention that we described in the previous section, targeting the misconception that empirical arguments are proofs of mathematical generalizations, was originally developed in a university-based design experiment that we conducted in a mathematics course for preservice elementary teachers in the United

States (Stylianides and Stylianides 2009). To use language from the previous quotation, this course was the setting where we framed the envisioned learning in the area of proof and the means of supporting it as a paradigm case of the broader phenomenon of what it means to help individuals overcome the given misconception. It was the explanatory theoretical framework that we developed as part of this research that allowed the successful adaptation of the intervention, which in turn led to similarly promising results when the intervention was used in a new context with high-attaining secondary students in England (Stylianides and Stylianides 2014b). The theoretical framework comprised multiple elements. One important element concerned the use of mathematics tasks and other instructional means to engineer cognitive conflicts for students thus supporting them, through a process of conceptual change, to move away from the misconception that empirical arguments are proofs. We elaborate on aspects of this element of the theoretical framework in our discussion of Characteristic 3.

Characteristic 2: A Narrow and Well-Defined Scope

There is no doubt that the overall development of students' learning of proof can be achieved only by the synergistic effect of many interventions that collectively address a wide range of learning goals (including key and persistent problems of students' learning) and, collectively, extend over a considerable period of time. Also, an intervention will likely not achieve its goals once and for all, but rather it will likely require follow up reinforcement or solidification. Notwithstanding these points, however, the issue arises as to what might constitute the *unit* of an intervention in a collection of interventions that together promote students' learning of proof. We consider important that the unit be as small as realistically possible, i.e., that each intervention that aims to address a key and persistent problem of students' learning in the area of proof (or in another area of mathematics) has a narrow and well-defined scope, which in turn can allow the intervention to have also a relatively short duration.

Characteristic 2 offers three major advantages. The first advantage is that the fulfillment of this characteristic can help with achieving also Characteristic 1 discussed earlier. Specifically, it is easier to theorize an intervention that has a narrow and well-defined scope, and possibly a short duration too, due to the relatively smaller number of factors involved compared to an intervention with a wider scope and a longer duration. In particular, an intervention with Characteristic 2 makes it less problematic for researchers to construct analytical hypotheses about the substantiated means that supported the emergence of specific patterns in students' learning during the classroom implementation of the intervention (Stylianides and Stylianides 2014a).

The second advantage relates to the issue of "scaling up" educational innovations (e.g., Cohen and Ball 2007). It is more practical for teachers to incorporate into their existing curricula an intervention that has Characteristic 2, for otherwise a major restructuring of the curricula could be required (Stylianides and Stylianides 2013). Also, if teachers decompose into smaller parts an intervention with a wide scope and a long duration that was designed as one whole entity, the smaller parts will likely not carry the theory- and research-informed features that characterized the overarching intervention thus endangering its effectiveness. This could happen even when a teacher implemented all the parts but in a different order than in the overarching intervention.

The third advantage relates to the issue of the fidelity of implementation of an educational innovation. It is more likely that teachers will implement with higher fidelity an intervention that has Characteristic 2 than an intervention with a wider scope and a longer duration, for the latter kind of intervention would include more action points in the implementation plan from which a potential deviation could compromise the outcomes. Indeed, high fidelity of implementation of an intervention is important so as to preserve the theoretically essential components of the intervention and increase the likelihood of obtaining the expected outcomes (Yeager and Walton 2011).

Of course, it is one thing to desire an intervention with Characteristic 2, given all of the aforementioned advantages, and another to actually develop such an intervention. Is it realistic to develop interventions that have Characteristic 2? In their review of randomized experiments on psychological interventions in education, Yeager and Walton (2011) established this possibility in the area of psychological interventions, some of which had remarkably short duration (20 min) but nevertheless significant and lasting effects on students' academic achievement. Our intervention-oriented research in the areas of proof (Stylianides and Stylianides 2009, 2014b) and problem solving (Stylianides and Stylianides 2014a) suggests that, despite the fundamental differences between the psychological interventions discussed in Yeager and Walton (2011) and the kind of classroom-based interventions we discuss herein, Characteristic 2 can also be achieved in mathematics education.

In more detail, the way in which the proof intervention discussed in Stylianides and Stylianides (2009, 2014b) fulfills Characteristic 2 can be easily inferred from our description of that intervention in the previous section. Important to note here is how we tried to evaluate the effect of this and other interventions in the larger collection of short-duration interventions that we designed and implemented in the university course that provided the context for our design experiment. In the absence of an experimental research design, the effect of each individual intervention could not be completely isolated from other potential influences in the course. Yet the effect of each intervention was evaluated in two important ways.

The first way was by comparison of findings related to the intervention across the five research cycles of the design experiment. From one research cycle to the next there were no major differences in the design of each intervention, and so we were able to evaluate the impact of the changes (improvements) we were introducing over time. The second was by triangulation of findings from multiple data sources within the same research cycle. The data in the last research cycle of the design experiment included the following: videos and field notes of the implementation of the interventions, students' written responses to specially designed prompts (see notion of "conceptual awareness pillars" in the next section) or other student data (e.g., students' responses to tasks), students' responses to a survey and a mathematics test at the beginning and at the end of the course, students' responses to particular questions in the homework assignments for the course, and individual interviews with the students at the end of the course.

The effect of an intervention with Characteristic 2, including the role played by key factors in it, may also be examined with a randomized controlled trial. If the intervention can be used stand-alone and has a short duration, this can ease the handling of the numerous practical challenges. See, for example, the randomized controlled trials conducted by Jones et al. (2016): the interventions comprised only three lessons each and their versions in the different conditions varied by only one key factor, which corresponded to whether it is better to teach mathematical topics using abstract or contextualized representations.

Characteristic 3: An Appropriate Mechanism to Trigger and Support Conceptual Change

Our focus on key and persistent problems of students' learning in the area of proof (but also more broadly), such as common and hard-to-remediate misconceptions, implies another essential characteristic of classroom-based interventions that would aim to address these problems: the inclusion in the interventions of an appropriate mechanism that can help trigger and support conceptual change in students so that their new conceptions will align better with conventional knowledge. While the inclusion of one such mechanism is, we argue, an essential characteristic, the precise nature of the mechanism can vary according to the needs of each particular intervention and the learning goal it aims to address.

In the context of our design experiment that we described earlier, the notion of *cognitive conflict* (Piaget 1985) was at the core of the mechanism we had developed and refined over the years of our study to trigger and support conceptual change in our students. In the early stages of our study, we tried to engineer cognitive conflicts for our students by strategically bringing them against mathematical situations that contradicted their existing conceptions. Yet our experience in the study, which agreed with prior research and practice, showed that students tended not to recognize contradictions as problematic for their current conceptions, thus treating contradictions as exceptions without experiencing a cognitive conflict or feeling an "intellectual need" (Harel 1998) to modify their existing conceptions.

This problem served as a driving force for the development of our instructional design and resulted in the genesis of the notion of *conceptual awareness pillars* (Stylianides and Stylianides 2009) or simply *pillars*: these are instructional activities that aim to direct students' attention to key issues in a classroom situation (such as issues related to problematic aspects of students' learning targeted by an intervention), with a consequential (potential) increase in students' awareness of their conceptions about those issues. The notion of pillars embodies a relationship between attention and awareness whereby "[b]eing aware is a state in which attention is directed to whatever it is that one is aware of" (Mason 1998, p. 254). A pillar can take different forms: it can be a teacher's question for students to consider or reflect on an issue that was raised, or was intended to be raised, as a result of students' engagement with a task that aimed to provoke cognitive conflict for students; it can be simulated student talk or dialogue that raises a particular issue and creates a context for productive discussion about or reflection on the issue among students; and so on.

We found that having appropriately designed pillars prior to and after each "potential cognitive conflict" (Zazkis and Chernoff 2008) in our interventions increased the likelihood of these becoming actual cognitive conflicts for students, with more students engaging in reflection on the emerging contradictions in their conceptions. Specifically, a well-designed pillar before a contradiction (in the form of a counterexample) that was intended to create a cognitive conflict for students often directed students' attention to their current conceptions thus helping them become more aware of these conceptions, while a pillar after such a contradiction directed students' attention to the problematic nature of their original conceptions thus triggering among them a process of reflection and revision of these conceptions (Stylianides and Stylianides 2009, 2014a, b). Designing the pillars to be successful in serving these important functions was an act of "empirical tinkering" (Morris and Hiebert 2011) over the cycles of our design experiment. Important to note also is that the pillars were only one among several elements of the overall instructional design and underpinning theoretical framework that we used to trigger and support conceptual change among students in the area of proof; elaborating on other elements is beyond the scope of this chapter.

Conclusion

Classroom-based interventions with the three characteristics we discussed in this chapter are not common, but this is not surprising given that intervention-oriented research is not as developed as other kinds of research in the area of proof (Stylianides et al. 2016, 2017) but also more broadly (Stylianides and Stylianides 2013). As we explained earlier, we are not suggesting that interventions that do not have the three characteristics are lacking. After all, our focus has been on a particular kind of interventions: those that aim to address key and persistent problems of students' learning. Yet, we argued, interventions with the three characteristics have something important to offer in our efforts as a field to advance students' learning of proof (as well as learning in other areas of mathematics) and, thus, deserve more attention by future research.

If these characteristics guide the design of a collection of classroom-based interventions as part of a larger research program targeting key and persistent problems of students' learning in the area of proof, then at least one other characteristic will emerge as being crucially important: a coherent and articulated conceptualization of the nature of proof in school (or even university) mathematics. For example, our previous discussion of the misconception that a few confirming cases suffice to establish the truth of a mathematical generalization, in the intervention reported in Stylianides and Stylianides (2009), implies a conceptualization of proof whereby empirical arguments are excluded from the set of arguments that meet the standard of proof. The many benefits of having a clear conceptualization of the nature of proof have already been discussed in the literature (e.g., Balacheff 2002; Reid 2005; Stylianides 2007, 2016a). Beyond these benefits, the possible lack of such a conceptualization can create serious problems (pedagogical, epistemological, theoretical) in future efforts to put together existing interventions with the three characteristics in a comprehensive curriculum for school (or university) mathematics to address key and persistent problems of students' learning in the area of proof. For example, an inconsistent meaning of proof across these interventions can exacerbate the problem of diversity of teachers' perceptions of proof in the curriculum and of the different pedagogical approaches to proof one might expect to see implemented in these teachers' classrooms (Davis, this volume).

A related priority for research in this area is to find ways to make the research knowledge produced by promising classroom-based interventions accessible to, and usable by, many teachers. Without underestimating the complexity of the problem, a possible way to make progress in addressing it would be through curricular resources that aim to support teacher learning alongside student learning of proof. Such curricular resources—often referred to as *educative curricular resources*—have the potential to support instructional reform (e.g., Ball and Cohen 1996) and are particularly important in the area of proof given that not only students but also many teachers face difficulties with proof (e.g., Harel and Sowder 2007; Stylianides et al. 2017). Unfortunately, though, the treatment of proof in existing curricular resources (notably textbooks) is limited in terms of supporting student or teacher learning. This has been indicated, for example, by the results of a collection of studies published in a special issue (Stylianides 2014) and elsewhere (Davis 2012; Sears and Chávez 2014; Stylianides 2008, 2009; Thompson et al. 2012; Wong and Sutherland, this volume). Obviously, there is a long way to go before it becomes the norm for curricular resources to include rich and appropriately designed tasks for students to engage with proof as well as detailed instructional plans and guidance for teachers for how they may implement promising interventions in the classroom. Toward this end, we (as a field) need appropriate frameworks for the design of different kinds of tasks that can afford students with the range of opportunities required for broad learning in the area of proof (for discussion of one such framework, see Stylianides 2016a). Also we need appropriate frameworks and analytic approaches that can allow us to study, in connected ways, the affordances of the tasks in the interventions and factors that influence their classroom implementation (for discussion of such frameworks and approaches, see Stylianides 2016b).

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.

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

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

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

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

Cohen, D. K., & Ball, D. L. (2007). Innovation and the problem of scale. In B. Schneider & S. McDonald (Eds.), *Scale-up in education: Ideas in principle* (Vol. I, pp. 19–36). Lanham, MD: Rowman & Littlefield.

Davis, J. D. (2012). An examination of reasoning and proof opportunities in three differently organized secondary mathematics textbook units. *Mathematics Education Research Journal, 24,* 467–491.

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 32*(1), 5–8.

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

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

Greeno, J. G. (2006). Theoretical and practical advances through research on learning. In J. L. Green, G. Camilli, & P. B. Elmore (with A. Skukauskaite & E. Grace) (Eds.), *Handbook of complementary methods in education research* (pp. 795–822). Washington, D.C.: American Educational Research Association.

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

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

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

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

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

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

Jones, I., Inglis, M., Gilmore, C., & Bisson, M. J. (2016). *Measuring conceptual understanding: The case of teaching with abstract and contextualised representations* (Final Project Report). London: Nuffield Foundation. Retrieved November 10, 2016 from http://www.nuffieldfoundation.org/sites/default/files/files/MCU_FINALREPORT.pdf.

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

Mason, J. (1998). Enabling teachers to be real teachers: Necessary levels of awareness and structure of attention. *Journal of Mathematics Teacher Education, 1,* 243–267.

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

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

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

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

Reid, D. (2005). The meaning of proof in mathematics education. In M. Bosch (Ed.), *Proceedings of the 4th Conference of the European Society for Research in Mathematics Education* (pp. 458–468). Sant Feliu de Guixols, Spain. Retrieved December 11, 2011 from http://ermeweb.free.fr/CERME4/CERME4_WG4.pdf.

Ruthven, K., & Goodchild, S. (2008). Linking researching and teaching: Towards synergy of scholarly and craft knowledge. In L. D. English (Ed.), *Handbook of international research in mathematics education* (2nd ed., pp. 561–588). New York, NY: Routledge.

Sears, R., & Chávez, O. (2014). Opportunities to engage with proof: The nature of proof tasks in two geometry textbooks and its influence on enacted lessons. *ZDM—The International Journal on Mathematics Education, 46*(5), 767–780.

Stevenson, A., & Lindberg, C. A. (Eds.). (2012). *New Oxford American Dictionary* (3rd ed.). Oxford: Oxford University Press. http://www.oxfordreference.com/view/10.1093/acref/9780195392883.001.0001/acref-9780195392883. Accessed March 25, 2013.

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

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

Stylianides, A. J., Bieda, K. N., & Morselli, F. (2016). Proof and argumentation in mathematics education research. In A. Gutiérrez, G. C. Leder, & P. Boero (Eds.), *The second handbook of research on the psychology of mathematics education* (pp. 315–351). Rotterdam, The Netherlands: Sense Publishers.

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

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

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

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

Stylianides, G. J. (Ed.). (2014). Reasoning-and-proving in mathematics textbooks: From the elementary to the university level. *International Journal of Educational Research* [Special Issue], *64,* 63–148.

Stylianides, G. J. (2016b). *Curricular resources and classroom use: The case of mathematics*. New York: Oxford University Press.

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

Stylianides, G. J., & Stylianides, A. J. (2014b). The role of instructional engineering in reducing the uncertainties of ambitious teaching. *Cognition and Instruction, 32*(4), 374–415.

Stylianides, G. J., & Stylianides, A. J. (Eds.). (2017). Research-based interventions in the area of proof. *Educational Studies in Mathematics* [Special Issue], *96*(2), 119–274.

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

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

Wiliam, D., & Lester, F. K. (2008). On the purpose of mathematics education research: Making productive contributions to policy and practice. In L. D. English (Ed.), *Handbook of international research in mathematics education* (2nd ed., pp. 32–48). New York, NY: Routledge.

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

Zaslavsky, O., & Ron, G. (1998). Students' understanding of the role of counter-examples. In A. Olivier & K. Newstead (Eds.), *Proceedings of the 22nd Annual Meeting of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 225–232). Stellenbosch, South Africa.

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

Chapter 8
How Can a Teacher Support Students in Constructing a Proof?

Bettina Pedemonte

Abstract This chapter analyzes the one-to-one interaction between student and teacher when student is engaged in constructing a geometrical proof. This analysis shows that it is not easy for the teacher to modify the student's argumentation based on conceptions that can hardly evolve into theorems. The teacher's intervention can be considered effective if it doesn't completely "interrupt" cognitive unity between the student's argumentation and proof, but opposite it encourages the continuity between them. Toulmin's model, used to analyze the student's argumentation and the teacher's intervention, highlights that the teacher's intervention needs to become a rebuttal in student's argument to invalidate it. The incorrect argument is refused by student only if the teacher's rebuttal has the same backing of the student argument and it is "coherent" with the student warrant.

Keywords Argumentation · Proof · Cognitive unity · Conceptions
Toulmin's model · Teaching

Introduction

Proving is tightly connected to the on-going argumentation activity involved in solving a problem (Boero et al. 1996; Pedemonte 2005, 2007). Engaged in mathematical problem-solving, learners proceed based on their understanding of mathematical concepts and related process. This activity doesn't initially have the structure of what will be considered as mathematical proof. Instead, it is a tangle of intuitions, know-how, knowledge and a variety of mental constructs allowing learners to make choices and to take decisions. Following Confrey (1990) and Balacheff (2009), I use the term *conception* to refer to these complex mental structures. When students construct their argumentations to produce proof, they use their conceptions (Balacheff 2009), which are at the basis of argumentation activity,

B. Pedemonte (✉)
San Jose State University, San Jose, USA
e-mail: bettina.pedemonte@sjsu.edu

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_8

even if in the proof (considered as final product in the proving activity) they might not be present. Educational research analyzed how students' conceptions strongly affect the construction of proof (Pedemonte 2005) and how different students' conceptions interact in the construction of a single proof (Pedemonte and Balacheff 2016). This research showed that sometimes students are not able to construct proof because argumentations are based on conceptions that cannot evolve into theorems. From a didactical point of view, it is important to find solutions to help students to overcome this difficulty. This chapter can be considered an extension of these studies. It focuses on interaction between teacher and students when the students are constructing proof to solve an open problem (Arsac 1991) in Geometry. The aim of this chapter is to analyze which kind of teacher intervention can be considered effective to support students in constructing proof when they are mobilizing conceptions that hardly can evolve into theorems.

Cognitive Continuity Between Argumentation and Proof

Educational research highlighted that when students solve an open problem using an argumentation activity to construct a conjecture, continuity between argumentation and proof, called *cognitive unity* (Boero et al. 1996) can be observed.

> During the production of the conjecture, the student progressively works out his/her statement through an intensive argumentative activity functionally intermingled with the justification of the plausibility of his/her choices. During the subsequent statement-proving stage, the student links up with this process in a coherent way, organizing some of previously produced arguments according to a logical chain. (Boero et al. 1996, p. 119)

This phenomenon is referred to by the authors as *cognitive unity*. Experimental research about cognitive unity (Boero et al. 1996; Garuti et al. 1996, 1998; Pedemonte 2005, 2008) shows that proof is more achievable to students if an argumentation activity is developed for the construction of a conjecture. Indeed, students can construct the proof by organizing some of the previously produced arguments in a logical chain.

When students solve open problems they generally produce an argumentation to construct and/or justify their conjecture. This argumentation can be used to construct a proof validating the conjecture. Thus, the relationship between argumentation and proof is connected to the relationship between conjecture and theorem. A theorem is composed of three elements: a statement, a proof and a mathematical theory (Mariotti et al. 1997). The theorem exists because there is a mathematical theory (a system of shared principles and deduction rules) which allows construction of a proof, thereby validating the statement. Likewise, a conjecture is constituted by a statement, an argumentation and a system of conceptions (Pedemonte 2005). The argumentation can be related to the conjecture in two ways: it can contribute to the construction of a conjecture, so it precedes the statement, or it can justify a conjecture, previously constructed as a "fact", so it comes afterwards. In both cases, conjecture is based on arguer's conceptions (Balacheff 2009). These conceptions belong to the arguer's system of knowledge, that is not necessarily a mathematical theory.

Therefore, some conceptions can prevent *cognitive continuity* (Garuti et al. 1996; Pedemonte 2005) between argumentation and proof. I refer to *incorrect conceptions*, namely conceptions that cannot evolve into theorems because they are not supported by a mathematical theory. When the argumentation supporting the conjecture is based on incorrect conceptions, two possibilities can be envisioned: (1) the proof is not constructed because the student cannot replace the incorrect conception by a theorem (2) an "incorrect proof" is produced based on the conception used in the argumentation. In both cases, students need to change the resolution strategy to construct a proof. A didactical intervention could be very useful to help student to invalidate the incorrect conception to construct a different argument supporting a new conjecture. However, from a didactical point of view, it is a challenge to understand which intervention could be effective to support the student in changing the strategy to solve the problem in a different way. As highlighted by Tsujiyama and Yui (this volume) examples of unsuccessful arguments can facilitate students' reflection on their process of planning a proof and also this aspect should be account by the teacher.

This chapter shows that a teacher's intervention can be considered effective if it doesn't completely "interrupt" *cognitive unity* between the student's argumentation and proof, but opposite it encourages the continuity between them. The teacher should not "replace" the incorrect student's conception used to construct the conjecture, with a new one. The teacher's intervention needs to be part of the student's argumentation to maintain the continuity between student's argumentation and proof.

Toulmin's model is used to analyze how teacher's intervention can be part of student's argument (as rebuttal) and under which conditions it becomes a new warrant in the student's argument. I will show that the teacher's intervention seems to be effective if it "acts' on the warrant and the backing of the student's argument.

Toulmin's Model

In this chapter, Toulmin's model is used to analyze when teacher's intervention can support *cognitive unity* between the student's argumentation and proof. In particular, the model is used here not to focus in the relationships that exists between arguments, but in the relationships between the different components of a given argument. This is also the reason why this model was embraced by a large number of researchers in mathematics education (Inglis et al. 2007; Knipping 2008; Krummehuer 1995; Lavy 2006; Nardi et al. 2012; Pedemonte 2005, 2007, 2008; Weber and Alcock 2005; Wood 1999; Yackel 2001; Yackel and Rasmussen 2002).

An argument provides a standpoint (an assertion, an opinion) which is called a *claim* in Toulmin's terminology. *Data* are produced supporting the claim. A *warrant* provides the justification for using the data in support of the data-claim relationships; it can be expressed as a principle or a rule and it acts as a bridge between the data and the claim. This is the ternary base structure of an argument, but auxiliary elements may be necessary to describe it. Toulmin describes three of them: the *qualifier*, the *rebuttal* and the *backing*. The warrant imparts different degrees of force to the conclusion it justifies, which may be indicated by a qualifier such as 'necessarily', 'probably' or 'presumably' attached to the transition from the data to the claim. In the latter case, we may need to mention conditions of rebuttal "indicating circumstances in which the authority of the warrant would have to be set aside" (Toulmin 1958, p. 101). So, a warrant can be defended by appeal to a backing that can be expressed in the form of categorical statements of fact (Toulmin 1958, p. 105). A backing can be provided by a system of taxonomic classification, by a statute, by statistical results, or by a mathematical theory. The type of the backing could change greatly as one moves from one field of argument to another (Toulmin 1958, p. 104). Then, Toulmin's model of argument contains six related elements organized as showed in Fig. 8.1.

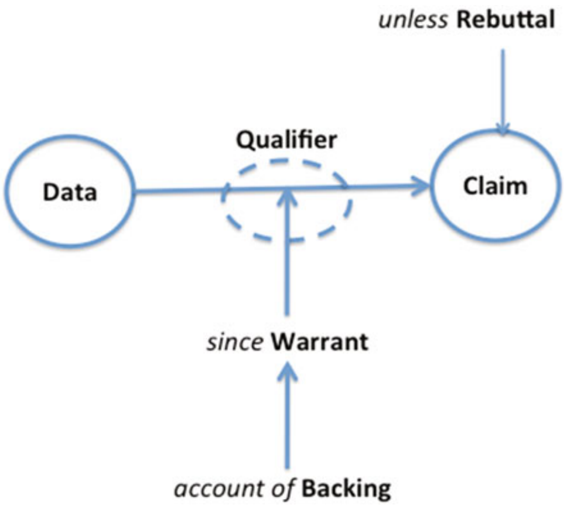


Fig. 8.1 Toulmin's model of argumentation

> Some warrants authorize us to accept claim unequivocally, given the appropriate data - these warrants entitle us in suitable cases to qualify our conclusions with the adverb "necessarily"; other authorize us to make the step from data to conclusion either tentatively, or else subject to conditions, exceptions or qualifications. (Toulmin 1958, p. 100)

In this last case, warrants are not based on theorems but on student's conceptions. These conceptions (that can appear explicit in the warrant or in the backing of the student's argument) do not necessarily lead to correct conclusions (Pedemonte and Balacheff 2016). When this is the case I can consider them as *incorrect conceptions*. When student uses an incorrect conception to construct a conjecture, the claim of the argument is in general not correct. Therefore, a teacher's intervention could be necessary to invalidate the argument and help student to modify the resolution strategy to solve the problem.

Thus, the teacher's intervention assumes the role of rebuttal in student's argumentation. Indeed, a statement can be considered as a rebuttal in the argument when it is the basis for incorporating into dialectical interaction a role of the opponent (Freeman 1991; Slob 2006). Toulmin's rebuttal allows for the introduction of counter-considerations provided by the *voice of the other* (Slob 2006). The teacher's intervention as *voice of the other* can be incorporated as rebuttal in the student's argument.

In the next section, two examples of teacher's intervention are provided to see how they can be incorporated in the student's argumentation.

Method

The following case studies are taken from a research project designed to analyze how teacher interventions affect students in the construction of proof. I have observed that when teacher interacts with student during a problem-solving activity, the student's argumentation can be strongly affected by the teacher intervention.

I have analyzed sixteen 11th grade Italian students interacting with their teacher while they were solving the following geometrical open-problem (taken from Camargo et al. 2007).

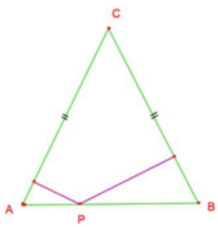


Fig. 8.2 Figure presented in Cabri-Geometry

In isosceles triangle ABC, determine the position of the point P, on the base of the triangle, so that the sum of the distances from P to the congruent sides of the triangle is minimum. Justify your answer.

This is not a standard problem for the analyzed students. However, students' background knowledge and skills were wide enough to allow them to construct a conjecture and to produce a proof. At the mentioned grades, they have been taught mathematical proof, hence they were aware that the current didactical contract required the construction of a proof even though this is not written in the description of the task. They knew all theorems necessary to solve the proposed problem. Furthermore, they were familiar with the use of Cabri-Geometry where the problem was presented. Students were also provided with a paper with the statement of the problem and room to write comments, calculations and responses to the questions.

The experiment lasted about 30 min.

In general, the students started the exploration moving the point P on the base of the triangle. Some students explored the "limit case" locating P on one of the endpoints of the segment AB. Some students inserted other points on the segment AB and drew auxiliary lines to have the possibility to compare different distances. In general, it was not spontaneous for students to realize that the sum of distances is invariant. The teacher's intervention was often necessary to help students to find a correct strategy to solve the problem.

The analysis of the relationship between the student's argumentation and the teacher's intervention is based on Toulmin's model. This scheme is used to analyze whether the teacher's intervention can be enclosed in student argumentation, which role it assumes inside it and under what conditions it is strong enough to affect the student's argumentation and proof.

Analysis

It was observed that the teacher's didactical intervention was often not effective in supporting student in the construction of the proof.

Two interactions between student and teacher have been selected from a case study. The purpose of the analysis is to highlight two different ways students' argumentations can be developed when a teacher's intervention is constructed to support students in the construction of a proof:

- Case 1: the teacher's intervention doesn't affect the student's proof
- Case 2: the teacher's intervention runs as an external rebuttal in the student's argumentation and it is useful to the student to construct proof.

Students' and teacher's argumentations were audio recorded and transcribed. Their utterances were selected and the argumentative steps were reconstructed. The texts have been translated from Italian into English.

An extract of Case 1 and Case 2 are presented below.

The student's text is in the left column while comments and analyses are reported in the right column.

Case 1. Valentina is exploring the problem in Cabri-Geometry. She moves point P on the segment AB silently. Then, she inserts another point in the segment AB and she seems to compare the distances: point P is situated in the middle of the segment AB, the other point is moved on AB (Fig. 8.3).

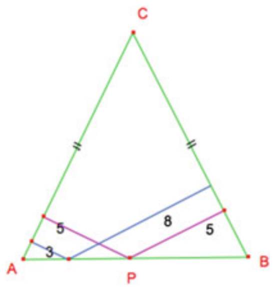
<table>
<tr>
<td>
1. V: …According to me the sum of the distances is minimum when… probably, when point P is in the middle
2. T: Why?
3. V: Because if you insert point P here, this distance (shows segment 8) is longer (in respect to segment 5)
4. T: Yes, but the other distance (segment 3) is shorter…
5. V: Yes, but less, I mean… for example if this distance is 5 and the other is 5… in this case, this distance could be 8 and this distance could be 3… however longer than the other.
</td>
<td>


Fig. 8.3 Figure constructed by Valentina in Cabri-Geometry

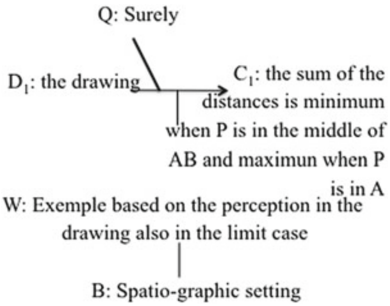
The distances (5, 5 and 3, 8) are not really calculated. Valentina suggests these distance to support her conjecture but she doesn't verify in Cabri-Geometry. Valentina's argument can be represented as follows:

D_1: the drawing → C_1: the sum of the distances is minimum when P is in the middle of AB
W: Exemple based on the perception in the drawing
B: Spatio-graphic setting
</td>
</tr>
</table>

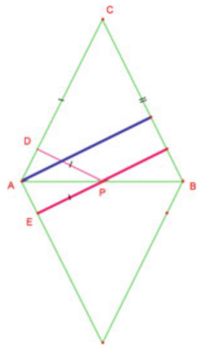
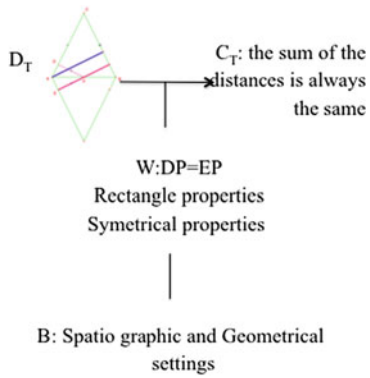
At this point, the argumentation starts and Valentina constructs a conjecture. She thinks that the sum of distances is minimum when point P is situated in the middle of segment AB.

Let's see the interaction between Valentina and the teacher.

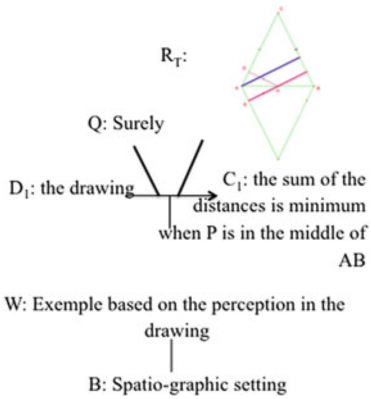
The teacher tries to modify Valentina's conjecture acting on the qualifier of her argument. She asks Valentina if she is sure about her reasoning. However, despite teacher intervention, Valentina doesn't change her argument. She considers the limit case to validate her conjecture.

6. T: Are you sure? 7. V: Yes, if you consider the limit case *(the point is moved on point A)* … this is the longest *(the segment 8, that represents now the limit case)*, while P is in the middle you have the minimum. 8. T: mmm, ok…	*Valentina argument can be represented as follows:* 

The teacher proposes a new figure to help Valentina to modify her conjecture: in Cabri-Geometry she constructs the symmetrical triangle in respect to segment AB. The reflecting image is congruent to the original triangle. Then she constructs the symmetrical of segment DP in respect to the segment AB (as shown in Fig. 8.4). The idea of the teacher is to help Valentina to see that the figure constructed inside the two triangles is a rectangle. In this way, Valentina could deduce that the sides of a rectangle are equal and then the two distances are equal.

8. T: … but consider these two paths… you can construct the symmetrical triangle. You can see that the reflected image is congruent to the triangle ABC. Now, we can construct the segment that is symmetrical to segment DP (*she constructed segment EP*). Observe this quadrilateral. What can you see? 9. V: What can I see? 10. T: For example, look at these two sides of the quadrilateral … 11. V: I see that this side is equal to this side *(Valentina is considering the rectangle)* 12. T: And what about the angles? 13. V: They are right angles 14. T: So this quadrilateral is a… 15. V: A rectangle?	 Fig. 8.4 Figure constructed by the teacher to help Valentina to modify her argument *We can represent the teacher argument in the following way:*
	

The teacher's intervention is a rebuttal in Valentina's argument because it introduces counter-considerations to invalidate Valentina's conjecture and to help her modify her resolution strategy. However, this rebuttal is not strong enough to modify Valentina's conjecture. As a matter of fact, when the teacher asks Valentina what is the minimal distance, her answer is still the same: the minimal distance is when P is in the middle of the segment AB.

16. T : Yes! It is a rectangle… this side is equal to this side. So what is the minimal distance? 17. V: The minimal distance? It is … the minimal distance is when point P is in the middle!	*Valentina argument can be represented as follows:* 

Despite the teacher intervention Valentina does not solve the problem and does not change her conjecture. The reason is that the teacher's intervention does not affect Valentina's conception.

If we compare the teacher's argument and Valentina's argument we can see that backings are different in the two arguments: Valentina is looking at the drawing, the backing of her argument is in the spatio-graphic setting. The teacher argument is in the spatio-graphic setting and in the geometrical setting. There is no connection between the two conceptions even if Valentina follows the teacher's reasoning.

The rebuttal provided by the teacher is not considered by Valentina as a counter-argument but as a different argument, not connected to her reasoning. For the teacher, the symmetrical triangle is useful to prove the equality among distances, but Valentina does not understand it. She does not see the connection between the symmetrical triangle and the problem.

In this example, the teacher's intervention seems to "interrupt" *cognitive unity* between the student's argumentation and proof. However, Valentina refuses the teacher's intervention and she reconstructs the continuity in her argumentation following her own reasoning based on her previous conception.

Valentina does not construct proof to the problem.

Case 2. Francesco is exploring the problem in Cabri-Geometry. He moves point P on the base AB until it coincides with point A. He constructs the segment AD (as shown in Fig. 8.5) to represent the limit case. Then he moves point P in another position along segment AB to compare segment AD with the sum of the distances of point P from the two sides of the triangle: EP + PF in Fig. 8.5. Francesco thinks that the sum EP + PF is longer than segment AD.

1.F: I do not know…I really do not know
2.T: Try to construct a conjecture
3.F: it seems to me that we can obtain the minimum sum when P is in A… or in B… Is it correct?
4.T: Let's see… Why according to you this segment AD is shorter than the sum EP +PF?
5.F: because you see… it seems to me that to do EP+PF you have to do more. AD is a straight line

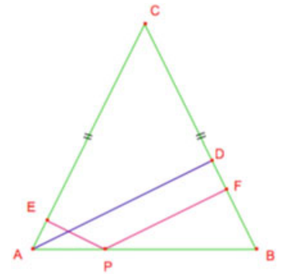
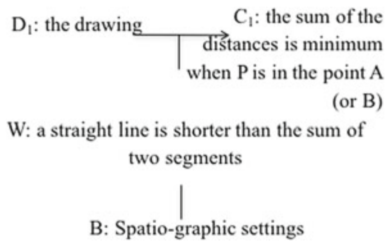


Fig. 8.5 Figure constructed by Francesco in Cabri-Geometry

Francesco argument can be represented in the following way:



To help Francesco, the teacher suggests him to construct a segment perpendicular to AD. This perpendicular segment is useful to make visible that segment AD is equal to the sum of the segments EP and PF. Starting from Francesco drawing, the teacher asks the student to add an "element", the segment PQ (as shown in Fig. 8.6), that can help Francesco to modify his reasoning, (not to produce another reasoning as in Case 1).

The teacher is analyzing the drawing from a geometrical point of view: RD is equal to PF because CB and QP are parallel as far as RD and PF. Furthermore, QP is perpendicular to RD. So the figure RPFD is a rectangle. In a similar way, AR and EP are equal because the two triangles AEP and ARP are congruent. The teacher is trying to help Francesco see that in the drawing segments EP and AR are equal as far as segments RD and PF. The new configuration that is constructed by the students under the guidance of the teacher offers a way to compare the two lengths.

The teacher is moving inside the spatio-graphic setting, where Francesco argument is constructed, but the backing of the teacher's argument is Geometry.

Therefore, she is connecting the spatio-graphic setting (backing in Francesco's argument) with the geometrical setting.

The teacher's intervention does not replace Francesco's conception but it does help him invalidate it acting in the warrant and in the backing of the student's argument.

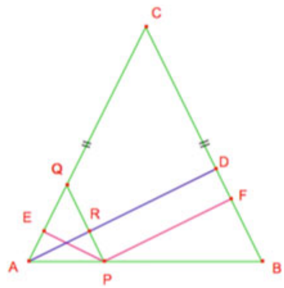
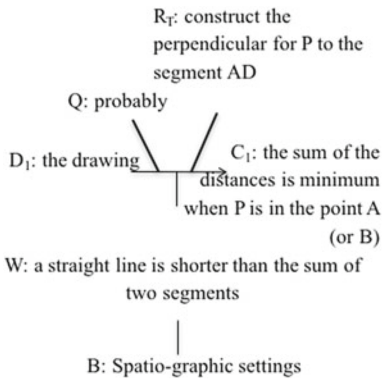
<table>
<tr>
<td>


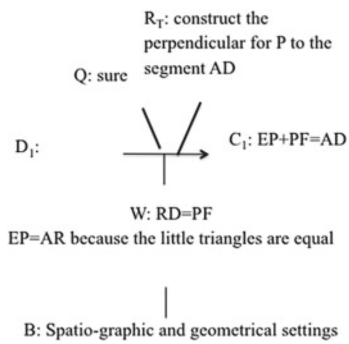
Fig. 8.6 Francesco constructs the segment AD and the teacher asks him to construct QP

6. T: Ok, what happen if I construct the perpendicular for P to this segment (*AD*)? Try to construct it.

Francesco constructs the perpendicular and labels the points Q and R.

7. F: RD is equal to PF and…oops…it seems that… no because this is
</td>
<td>
The teacher's intervention runs as a rebuttal in student's argument, as shown in the following representation.



The student's argument changes after the teacher's intervention.
</td>
</tr>
<tr>
<td>
longer…no wait these two segments are equal. I was wrong: EP+PF is equal to AD because EP is equal to AR because… These are two equals triangles!
</td>
<td>


</td>
</tr>
</table>

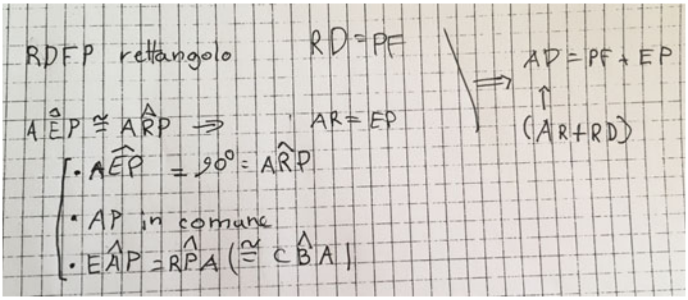


Fig. 8.7 Proof constructed by Francesco

In this second example the intervention of the teacher is useful to modify Francesco's argumentation. The rebuttal provided by the teacher is useful to modify Francesco argument because it can invalidate it, acting on the backing. The backings in Francesco's argumentation and in the teacher's argumentation are very closed. Even if the teacher's intervention is oriented in the geometrical setting, Francesco can follow it because the segment he inserts in his drawing runs as a link between the spatio-graphic and the geometrical settings.

Cognitive unity between student's argumentation and proof is not interrupted by the teacher's intervention. Opposite, it supports the continuity in Francesco's argumentation because it becomes part of it. Francesco solves the problem and constructs a proof as shown in Fig. 8.7.

Discussions and Conclusions

The use of Toulmin's model (1993) highlighted some important aspects. We have observed that there is a tight relation between conceptions summoned by student and the *modal qualifier* and the *rebuttal* in his argumentation (Toulmin 1958; Pedemonte 2005). If student conception is incorrect, an *external rebuttal* (produced by a person different from the arguer) can be constructed to invalidate the argument. We have observed that rebuttal produced from the teacher is effective when it acts on the warrant and the backing of the student's argumentation. The teacher's rebuttal should have the same backing of the student argumentation and it should be "coherent" with the student warrant (Case 2).

In other words, the *cognitive continuity* between argumentation and proof (Pedemonte 2005) should be maintained to make rebuttals effective for the construction of proof. In Case 2 the teacher's intervention is effective because it does not interrupt the *cognitive continuity* between the student's argumentation and proof. The teacher's intervention acts as a rebuttal in the student's argument invalidating his conception but not replacing it, as in Case 1. In Case 2, the teacher's intervention supports the cognitive continuity between argumentation and proof and the student can construct proof even if his initial conception was not correct. Opposite, in Case 1 the teacher is not able to invalidate the student's conception because she replaces the student's argument by her own argument. The external rebuttal is not effective to modify the student's argument because it does not affect the warrant and the backing of the student's argument.

This analysis was performed on a limited number of students, so results cannot be generalized. This is a work in progress research that needs to be experimented in other contexts and with a larger number of students.

However, even if the experiment was not part of the students' regular mathematics classes, this research has classroom-based issues related to proof and proving. The fine grain analysis of the teacher's intervention, and specifically the analysis of teacher–student interaction is of great interest for the development of a more complex analysis of classroom interaction on argumentation and proof. The analysis of the teacher's interventions in respect to the students' arguments shows how students arguments can be affected by the teacher intervention. This resonates with the results of Goizueta and Mariotti (this volume) that show how students exploit teacher's interventions to produce arguments to validate different mathematical models within a problem-solving situation. It is the quality of the intervention and the role that it assumes inside the specific student's argument that can modify the claim of the student. The teacher's intervention can run as a rebuttal inside a student's argument and not inside another one. Consequently, inside the class it is probably more difficult for the teacher to choose the "good intervention" to support students in the construction of proof.

References

Arsac, G., Germain, G., & Mante, M. (1991). *Problème ouvert et situation-problème*. Lyon: IREM.

Balacheff, N. (2009). Bridging knowing and proving in mathematics: A didactical perspective. In G. Hanna, H. N. Jahnke, & H. Pulte (Eds.), *Explanation and proof in mathematics. Philosophical and educational perspectives* (pp. 115–135). New York: Springer.

Boero, P., Garuti, R., & Mariotti, M. A. (1996). Some dynamic mental processes underlying producing and proving conjectures. In H. L. Chick & J. L. Vincent (Eds.), *Proceedings of the 20th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 121–128). Valencia: PME.

Camargo, L., Samper, C., & Perry, P. (2007). Cabri's role in the task of proving within the activity of building part of an axiomatic system. In D. Pitta-Pantazi & G. Philippou (Eds.), *Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education* (pp. 571–580). Larnaca, Cyprus: CERME.

Confrey, J. (1990). A review of the research on students' conceptions in mathematics, science, and programming. In C. Courtney (Ed.), *Review of research in education* (Vol. 16, pp. 3–56). American Educational Research Association.

Freeman, J. B. (1991). *Dialectics and the macrostructure of arguments: A theory of argument structure*. Berlin: Foris.

Garuti, R., Boero, P., & Lemut, E. (1998). Cognitive unity of theorems and difficulties of proof. In A. Olivier & K. Newstead (Eds.), *Proceedings of the 22nd Conference of the International*

Group for the Psychology of Mathematics Education (Vol. 2, pp. 345–352). Stellenbosch, South Africa: PME.

Garuti, R., Boero, P., Lemut, E., & Mariotti, M. A. (1996). Challenging the traditional school approach to theorems: A hypothesis about the cognitive unity of theorems. In H. L. Chick & J. L. Vincent (Eds.), *Proceedings of the 20th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 113–120). Valencia: PME.

Inglis, M., Mejia-Ramos, J. P., & Simpson, A. (2007). Modelling mathematical argumentation: The importance of qualification. *Educational Studies in Mathematics, 66,* 3–21.

Knipping, C. (2008). A method for revealing structures of argumentation in classroom proving processes. *ZDM Mathematics Education, 40*(3), 427–441.

Krummheuer, G. (1995). The ethnography of argumentation. In P. Cobb & H. Bauersfeld (Eds.), *The emergence of mathematical meaning: Interaction in classroom cultures* (pp. 229–269). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lavy, I. (2006). A case study of different types of arguments emerging from explorations in an interactive computerized environment. *The Journal of Mathematical Behavior, 25,* 153–169.

Mariotti, M. A., Bartolini Bussi, M. G., Boero, P., Ferri, F., & Garuti, M. R. (1997). Approaching Geometry theorems in contexts: From history and epistemology to cognition. In: E. Pehkonnen (Ed.), *Proceedings of the 21st Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 180–195). Lahti, Finland: PME.

Nardi, E., Biza, I., & Zachariades, T. (2012). "Warrant" revisited: Integrating mathematics teachers' pedagogical and epistemological considerations into Toulmin's model for argumentation. *Educational Studies in Mathematics, 79*(2), 157–173.

Pedemonte, B. (2005). Quelques outils pour l'analyse cognitive du rapport entre argumentation et demonstration. *Recherche en didactique des mathématiques, 25*(3), 313–348.

Pedemonte, B. (2007). How can the relationship between argumentation and proof be analysed? *Educational Studies in Mathematics, 66*(1), 23–41.

Pedemonte, B. (2008). Argumentation and algebraic proof. *ZDM Mathematics Education, 40*(3), 385–400.

Pedemonte, B., & Balacheff, N. (2016). Establishing links between conceptions, argumentation and proof through the ck¢-enriched Toulmin model. *Journal of Mathematical Behavior, 41,* 104–122.

Slob, W. H. (2006). The voice of the other: A dialogical-rhetorical understanding of opponent and Toulmin's Rebuttal. In D. Hitchcock & B. Verheij (Eds.), *Arguing on the Toulmin's model. New essays in argument analysis and evaluation* (pp. 147–162). New York: Springer.

Toulmin, S. E. (1958). *The use of arguments*. Cambridge: Cambridge University Press.

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

Wood, T. (1999). Creating a context for argument in mathematics class. *Journal for Research in Mathematics Education, 30*(2), 171–191.

Yackel, E. (2001). Explanation, Justification and argumentation in mathematics classrooms. In M. Van den Heuvel-Panhuizen (Ed.), *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 33–40). Utrecht, Olanda: PME.

Yackel, E., & Rasmussen, C. (2002). Beliefs and norms in the mathematics classroom. In G. Leder, E. Pehkonen, & G. Toerner (Eds.), *Beliefs: A hidden variable in mathematics education?* (pp. 313–330). Dordrecht, The Netherlands: Kluwer.

Chapter 9
Proof Validation and Modification by Example Generation: A Classroom-Based Intervention in Secondary School Geometry

Kotaro Komatsu, Tomoyuki Ishikawa and Akito Narazaki

Abstract Recent curriculum reforms underline mathematical activity related to proof validation, but few studies have explicitly addressed proof validation at the secondary school level. This chapter reports on our study of this issue. We suggest a specific kind of task for introducing proof validation in secondary school geometry and define the meanings of proof validation and proof modification in terms of Lakatos's notion of the local counterexample. We briefly report on a classroom-based intervention implemented using such tasks in a lower secondary school in Japan. We then analyze the results of a task-based questionnaire conducted after the intervention to investigate how well the students did in proof validation and modification. The analysis shows that student failure in proof validation arose mainly from their difficulty with producing diagrams that satisfied the condition of the proof problem.

Keywords Proof validation · Proof modification · Example generation
Local counterexample · Classroom-based intervention · Task design

Introduction

Proving is a fundamental activity in mathematical practice, and proof and proving are recognized as central to the substantial mathematical learning of all students (e.g., Stylianides et al. 2017). In our research, proofs refer to deductive arguments

K. Komatsu (✉)
Shinshu University, Nagano, Japan
e-mail: kkomatsu@shinshu-u.ac.jp

T. Ishikawa
Nagano Municipal High School, Nagano, Japan

A. Narazaki
Tojaku High School, Toyota, Japan

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_9

that show the paths from the suppositions of statements/conjectures to the conclusions by using statements that are already accepted as true. Although empirical arguments based on examples are not considered as proofs, example generation plays several roles in proof-related activity (e.g., Buchbinder, this volume; Reid and Vargas, this volume). Proof-related activity includes not only constructing proofs, but also making conjectures, and examining the truth of conjectures/statements and the validity of proofs even after the proofs are constructed. In particular, our study considers this proof-related activity from the fallibilism of mathematics (Lakatos 1976). As shown in Lakatos's description of the mathematical history of polyhedra and uniform convergence, example generation after proof construction is helpful in improving statements and proofs (Komatsu 2017).

One aspect of such proof-related activity is *proof validation* (Selden and Selden 1995, 2003), in which one examines arguments constructed as proofs to determine whether the arguments constitute legitimate proofs (a more elaborate definition of proof validation will be given later). Several studies on proof validation have been conducted in the field of mathematics education, giving many insights into how undergraduate students, trainee and in-service teachers, and professional mathematicians do proof validation (Alcock and Weber 2005; Inglis and Alcock 2012; Knuth 2002; Ko and Knuth 2013; Segal 1999; Selden and Selden 2003; Weber 2008, 2010).

However, there remain two issues that need to be addressed. First, few studies have investigated proof validation at the secondary school level (McCrone and Martin 2004; Reiss et al. 2001). Recent curriculumreforms have emphasized mathematical activity involving proof validation; for example, the Common Core State Standards Initiative (2010) in the United States lists the activity of constructing viable arguments and critiquing the reasoning of others as one of the standards for mathematical practice. The latest national curriculum in England also declares that, as part of mathematical reasoning, "pupils should be taught to […] assess the validity of an argument and the accuracy of a given way of presenting information" (Department for Education 2014, pp. 5–6). Given these requirements, proof validation should be introduced into secondary school mathematics. The studies by McCrone and Martin (2004) and by Reiss et al. (2001) are relevant in that they investigate whether secondary school students can discern the invalidity of circular arguments (in which conclusions are used as suppositions). For instance, McCrone and Martin (2004) surveyed 18 American high school students and showed that only 22% of the students correctly judged a circular argument invalid. However, because proof validation involves more than identifying circular arguments, it remains necessary to address other types of proof validation.

Second, participant activity in most of the previous studies ended at proof validation, and few studies have focused on how the participants might modify arguments that they judge to be invalid. Alcock and Weber (2005) conducted a related study in which they requested participating undergraduates to check an argument and modify it if needed. However, Alcock and Weber did not focus on the undergraduates' modifications, treating them as a relatively minor topic. In our view, proof modification is equally as important as validation from an educational perspective because it can provide students with the opportunity to revise their mathematical knowledge, skills and thinking, and repeatedly taking this opportunity may help foster a more reflective attitude in students.

In this chapter, we report on our research into these two issues. The structure of the chapter is as follows. First, we illustrate a specific kind of task for introducing proof validation and modification into secondary school geometry. Second, we give a short report of a classroom-based intervention implemented using such tasks at a lower secondary school in Japan. Third, we analyze a post hoc task-based questionnaire to investigate how well the students did in proof validation and modification. Finally, we conclude by discussing the results of this questionnaire, the implications for teaching, and possible directions for future research.

Proof Validation and Modification in Secondary School Geometry

To introduce proof validation into geometry classes at the secondary school level, we consider a specific kind of mathematical task: *proof problems with diagrams* (Komatsu 2017; Komatsu et al. 2014). Proof problems with diagrams are defined as tasks in which statements are described with reference to particular diagrams with labels, typically featuring one diagram per problem. An example is shown in Fig. 9.1, where the statement can be proved by showing ∠QAC = ∠PBD and ∠QCA = ∠PDB using the inscribed angle theorem.

We adopt the specific interpretation of these tasks in which the relevant statements are considered with respect to certain general classes to which the diagrams given in the tasks belong, rather than only for the given diagrams. For instance, we take it that the statement in Fig. 9.1 argues that triangle AQC is always similar to triangle BPD if points A and B are on circle O and the stated condition is satisfied (e.g. Fig. 9.2a, b). This interpretation enables students to use examples for proof validation by drawing diagrams that satisfy the conditions of statements and deciding whether the constructed proofs are applicable to these diagrams. In the above example, the proof for the diagram shown in Fig. 9.1 is invalidated by Fig. 9.2b, which rejects the use of the inscribed angle theorem as the reason for showing ∠QCA = ∠PDB because angle QCA is no longer an inscribed angle

As shown in the diagram, two circles O and O' intersect at points P and Q, and two points A and B are located on circle O. Draw line AP and let point C be the intersection point of the line and circle O'. Draw line BQ and let point D be the intersection point of the line and circle O'. Prove ΔAQC ~ ΔBPD.

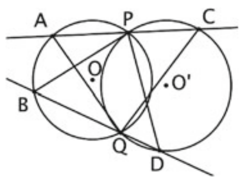


Fig. 9.1 Example of proof problem with diagram

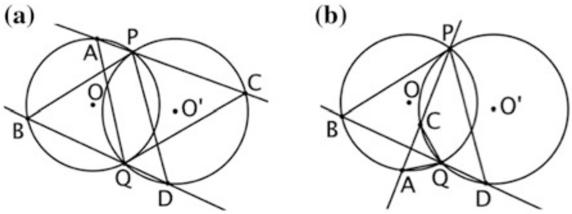


Fig. 9.2 Diagrams that satisfy the problem condition

corresponding to arc PQ of circle O′. The proof needs to be modified by altering this reason to the inscribed quadrilateral theorem, which states that 'an interior angle is equivalent to the exterior angle of the opposite angle.'

One of the characteristics of proof problems with diagrams is that the conditions of some (but not all) of the tasks are ambiguous because hidden assumptions may exist within the diagrams given in the tasks (Komatsu 2017; Stylianides 2007). As illustrated, these hidden assumptions can create opportunities for checking whether the constructed proofs are always valid by transforming the given diagrams. The tasks addressed in this chapter may appear as typical proof tasks in school geometry; students are usually presented with proof tasks that are described with diagrams and labels (Herbst and Brach 2006). However, our study capitalizes on the subset of proof problems with diagrams in which hidden assumptions exist in order to introduce proof validation into secondary school geometry.

To define proof validation and proof modification involving proof problems with diagrams, we borrow certain terms from Lakatos (1976), who differentiated between two types of counterexamples: *global counterexamples* and *local counterexamples*. A global counterexample, which is a counterexample in the conventional sense, refutes an entire statement, while a local counterexample rejects a step in a proof. In the aforementioned illustration, the case of Fig. 9.2b constitutes a local counterexample because it rejects the proof for the diagram shown in Fig. 9.1; however, it does not constitute a global counterexample because the statement, namely the similarity of triangles AQC and BPD, remains true. Based on Lakatos's terminology, we define proof validation as inspecting whether there are local counterexamples to proofs. Consequently, we define proof modification as, upon noticing the existence of local counterexamples, constructing proofs that are valid for the local counterexamples.

Note that the meaning of proof validation in this study is different from that in the literature. Previous studies on the subject have focused on identifying fallacies in invalid arguments, such as circular arguments, the falsity of reasons used to deduce certain lines, the mismatch of arguments in which the converses of target propositions are proved, and the inappropriateness of global counterexamples being used to show the falsity of propositions (e.g. Alcock and Weber 2005; Ko and Knuth 2013; McCrone and Martin 2004; Selden and Selden 2003). In contrast, our study deals with not-completely-invalid proofs, namely proofs that are valid in certain domains, and invites students to determine whether there are cases to which the proofs are not applicable.

By implementing and analyzing classroom-based intervention, we previously showed that individual tasks in the form of proof problems with diagrams are useful to some extent for eliciting mathematical activity relevant to proof validation and modification (Komatsu 2017). However, the students in our previous research were not fully involved in the proof validation because, in the implemented lessons, the local counterexample was given by the teacher, rather than being discovered by the students. This was because we anticipated that, as students are usually not familiar with proof validation, finding local counterexamples by transforming diagrams would be difficult for them. We address this issue in this study by designing task sequences through which students can be gradually introduced to proof validation and modification. In the following section, we briefly report one of the interventions implemented using the designed task sequence, and we analyze the results of our post hoc task-based questionnaire to investigate how well the students did in proof validation and modification.

Methods

Background

The study reported in this chapter was conducted as part of larger research on curriculum development for explorative proving (Miyazaki et al. 2016). The purpose of this study was to develop task sequences for facilitating student activity related to proof validation and modification in secondary school geometry through implementing classroom-based intervention. This chapter reports on one of the intervention, carried out by the second author of this chapter in his classroom at a lower secondary state school, involving 29 Japanese ninth-graders (aged 14–15 years old). The mathematical capabilities of the participating students were average for Japan according to the observation of the second author, who has 18 years of teaching experiences across several secondary schools in Japan.

Design of the Intervention

Three lessons (50 min per lesson) were used for our research: the first two lessons for the intervention, and the last lesson for the task-based questionnaire. The intervention was implemented over a relatively short duration (two lessons), with the specific purpose of introducing students to proof validation and modification (Stylianides and Stylianides, this volume).

The task sequence in the first two lessons was developed so that students could gradually experience proof validation and modification activity. For instance, because we anticipated that it would be difficult for students to discover local counterexamples independently by transforming diagrams without any prior experience, the first lesson was designed to present students with a diagram that constituted a local counterexample. Building on this experience, the second lesson was designed to invite students to discover local counterexamples by themselves. The proof problems used in the first and second lessons (which will be described in the results section) differed in terms of difficulty; in the second lesson, we used a task that was more difficult than the one in the first lesson.

The intervention was designed through close collaboration between the first author (researcher) and the second author (teacher). The first author initially drafted a task sequence and rough lesson plans, after which we discussed them in order to craft detailed lesson plans. In doing this, we took into consideration the usual practice in the classroom. Typical lessons in this classroom followed a problem-solving style consisting of task setup by the teacher, student individual and small-group work, and a concluding whole-class discussion (Stigler and Hiebert 1999). We organized our intervention in a similar way; the teacher began each lesson by providing one proof problem with a diagram, the students engaged in proof construction and then proof validation/modification, and finally the teacher led the whole class in a discussion where the students shared their thoughts regarding proof validation and modification. The second author carried out these lessons, and the first author observed them as a non-participant. We also held post-lesson discussions to reflect on the results of the lessons. During the lessons, the teacher used GeoGebra to demonstrate transforming diagrams, but the students were not allowed to use GeoGebra.

Task-Based Questionnaire

The third lesson was used for the task-based questionnaire (Fig. 9.3), which was based on the proof problem shown in Fig. 9.1. After proving the statement in Fig. 9.1, the students individually worked on this questionnaire for approximately 20 min. The students were intended to experience transforming diagrams in the first two lessons, and we expected that this experience could provide the students with the required mathematical background for considering different configurations of diagrams in the questionnaire. The students were familiar with the inscribed angle

Q1: Place point A on various places on circle O and find a case that rejects your proof.

Q2: Which part of your proof is rejected by this case?

Q3: Modify your proof to show $\Delta AQC \sim \Delta BPD$ even in this case.

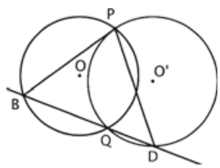


Fig. 9.3 Task-based questionnaire

theorem, the inscribed quadrilateral theorem, and proof construction for showing the similarity of triangles.

Students were regarded as successful in proof validation if they gave correct answers to Q1 and Q2. Students were regarded as successful in both proof validation and modification if they gave correct answers to Q1–3. We anticipated that there might be students who, in spite of producing diagrams that we as observers regard to be local counterexamples, were not able to answer which parts of their proofs were rejected by the diagrams; that is, they correctly answered Q1, but not Q2. These students were assessed as unsuccessful in proof validation because they drew the diagrams without recognizing them as local counterexamples.

Data Analysis

The classroom-based intervention was videotaped and transcribed, and the students' worksheets were collected. The following section gives a short report of the implemented intervention based on these data. After that, we focus on the students' answers to the task-based questionnaire. Based on the aforementioned criteria, the first and third authors independently assessed whether each student succeeded in proof validation and modification. We then synthesized our classifications, and any discrepancies were discussed until we reached a consensus. Afterwards, we categorized all the correct and incorrect answers to identify the main difficulties the students encountered. English translations of the original Japanese tasks and students' answers are given. All the students' names presented in this chapter are pseudonyms.

Results

The Implemented Intervention

The first lesson involved the task shown in Fig. 9.4, in which the students proved the statement by showing ∠BPA = ∠CPD and ∠BAP = ∠CDP using the equality of vertical angles and the inscribed angle theorem, respectively. The teacher then used GeoGebra to move point A and present the diagram shown in Fig. 9.4b, which the students recognized to be a local counterexample to their proofs because the previous reasons for ∠BPA = ∠CPD and ∠BAP = ∠CDP could not be employed. After that, the students engaged in proof modification by, for instance, changing the reason for ∠BPA = ∠CPD from the equality of vertical angles to the identity of the angles.

The second lesson involved the task in Fig. 9.5, in which the students proved the statement by showing ∠PAC = ∠PBD and ∠PCA = ∠PDB using the inscribed

As shown in the diagram, there are four points A, B, C, and D on circle O. Draw lines AC and BD, and let point P be the intersection point of the lines. Prove ΔPAB ~ ΔPDC.

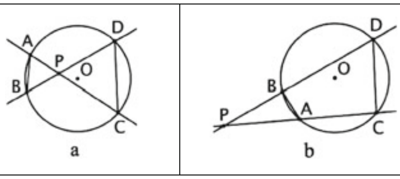


Fig. 9.4 Task and local counterexample shown by the teacher (the first lesson)

As shown in the diagram, two circles O and O' intersect at points P and Q, and two points A and B are located on circle O. Draw line AQ and let point C be the intersection point of the line and circle O'. Draw line BQ and let point D be the intersection point of the line and circle O'. Prove ΔPAC ~ ΔPBD.

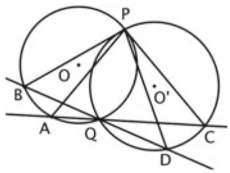


Fig. 9.5 Task used in the second lesson

angle theorem. Then, the teacher prompted the students to find local counterexamples to their proofs independently by drawing various diagrams. The local counterexample typically discovered by the students was the case shown in Fig. 9.6a, placing point A on arc PQ of circle O. They found that the reason for ∠PAC = ∠PBD did not apply, so they replaced it with the inscribed quadrilateral theorem. The teacher then introduced another local counterexample from a certain student's worksheet (Fig. 9.6b), where line AQ was a tangent line to circle O′ and points C and Q were regarded as being coincident. However, the class did not further examine this case because proof modification for it required another theorem, namely the alternate segment theorem, which the students had not yet learnt.

The third lesson involved the aforementioned task-based questionnaire. The students initially tackled the proof problem in Fig. 9.1. Because the teacher asked successful students to help students who found the proof difficult, all of the students completed full proofs. All of their proofs were based on showing ∠QAC = ∠PBD and ∠QCA = ∠PDB using the inscribed angle theorem. The following proof was written on the blackboard by a student, Takumi, and was shared in the classroom:

Regarding ΔAQC and ΔBPD,

Since inscribed angles corresponding to arc PQ in circle O are equal,

$$\angle QAC = \angle PBD \tag{9.1}$$

Similarly, since inscribed angles corresponding to arc PQ in circle O′ are also equal,

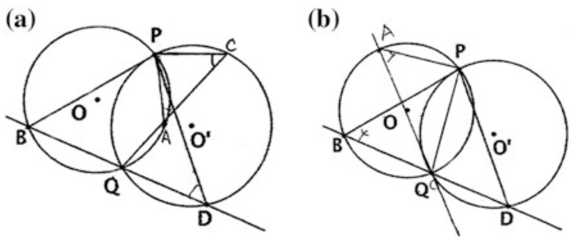


Fig. 9.6 Local counterexample in the second lesson (from the students' worksheets)

$$\angle QCA = \angle PDB \tag{9.2}$$

From (9.1) and (9.2), since the two pairs of angles are equal, ΔAQC ~ ΔBPD.

After that, the students individually worked on the task-based questionnaire. We failed to collect the questionnaires from two students, so the answers given by 27 students were analyzed.

The Results of the Task-Based Questionnaire on Proof Validation

Sixteen students (59%) succeeded in proof validation by producing diagrams that constituted local counterexamples to their proofs in Q1 and indicating which parts of their proofs were rejected by the diagrams in Q2. The local counterexample discovered by most students (12 students) was the case where point A was on arc PQ of circle O. For instance, Ren drew the diagram shown in Fig. 9.7a and answered in Q2 that "[the part stating] 'since inscribed angles corresponding to arc PQ in circle O are equal, ∠CAQ = ∠DBP' [is not valid]." Five students produced a case where point C was on arc PQ of circle O′ (Fig. 9.2b). The number of discovered local counterexamples was larger than the number of students who succeeded in proof validation because several students produced two or three local counterexamples.

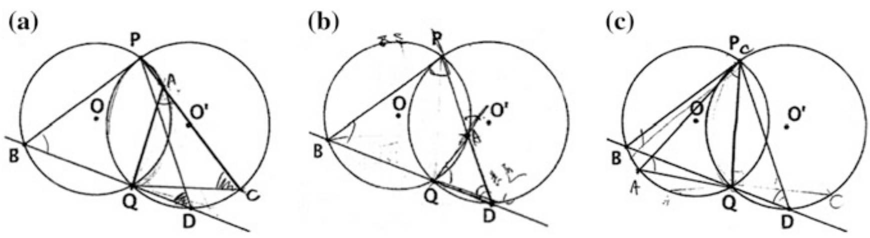


Fig. 9.7 Local counterexamples produced by students (from their worksheets)

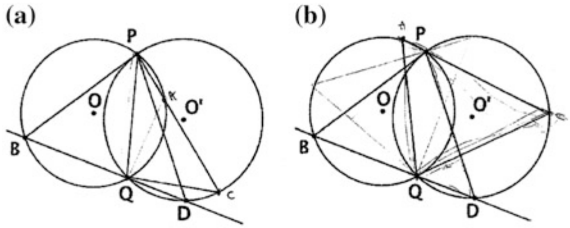


Fig. 9.8 Failure to discover local counterexamples (from the students' worksheets)

It is noteworthy that some students produced special local counterexamples. For instance, three of the above 12 students placed point A on the intersection point of segment DP and circle O (in this case, points C and D coincide); Shota drew the diagram shown in Fig. 9.7b and answered in Q2 that "it is not valid that the inscribed angles corresponding to circles O and O′ are equal." Four students drew a case where line AC was a tangent to circle O or O′ (Fig. 9.7c). The reasons why many students produced the case of Fig. 9.7a and several students produced special cases such as Fig. 9.7b, c would be related to the results of the implemented intervention, as such cases were shared with the whole class in the second lesson (Fig. 9.6).

There were 11 students (41%) who failed in proof validation. The most typical incorrect answer, given by four students, was derived from drawing diagrams that did *not* satisfy the condition of the problem. For instance, Misaki drew the diagram shown in Fig. 9.8a, making triangle PQC instead of triangle AQC. There were other incorrect answers, where three students drew diagrams that satisfied the problem condition but their proofs remained valid for these diagrams (Fig. 9.8b). These results indicate that finding local counterexamples by drawing diagrams was difficult for several students even after the intervention that explicitly focused on introducing them to proof validation and modification.

The Results of the Task-Based Questionnaire on Proof Modification

Ten students succeeded in proof modification, properly adjusting their proofs to cope with the local counterexamples they had discovered. This number accounts for 37% of all the students (10/27) and 63% of the students who succeeded in proof validation (10/16).

Kaito's answer to Q3 was typical for the successful students who considered the case of Fig. 9.7a:

> It is enough to change part (9.1) [of the initial proof] to 'since an interior angle and the exterior angle that is next to the opposite interior angle are equal in the quadrilateral inscribed to circle O, ∠PBD = ∠QAC.'

In this answer, he changed the reason for ∠PBD = ∠QAC from the inscribed angle theorem to the inscribed quadrilateral theorem. All of the students who considered the case of Fig. 9.7b performed the same modification, further changing the reason for the equality of another pair of angles, namely angles QCA and PDB, from the inscribed angle theorem to the identity of these angles.

Six students succeeded in proof validation, but failed in proof modification. Half of them attempted to modify their proofs by changing which pair of angles was referenced instead of the reasons for the equality of the angles. For instance, Aoi and Nanami considered the case of Fig. 9.7a and tried to modify the parts of their proofs showing ∠QAC = ∠PBD as follows:

Aoi: Since the four points P, B, Q, and A are inscribed to circle O, ∠BPA = ∠AQC.

Nanami: Since the quadrilateral PAQB is inscribed to circle O, ∠AQC = ∠BPD.

Aoi's answer is incorrect because angle BPA is equal with angle AQD, not angle AQC. Nanami's answer is insufficient because, while angle AQC is equal with angle BPD, this cannot be derived using only the inscribed quadrilateral theorem. It is necessary to show that ∠CQD = ∠CPD with the inscribed angle theorem, ∠AQD = ∠APB with the inscribed quadrilateral theorem, and then ∠AQC = ∠AQD − ∠CQD = ∠APB − ∠CPD = ∠BPD. However, she did not use this reasoning, only superficially changing the referenced pair of angles.

Proof modification for Fig. 9.7c requires the alternate segment theorem. One of the four students who produced this type of special case was assessed to be unsuccessful in proof modification because this was her only local counterexample, and she was not able to modify her proof due to a lack of knowledge of the alternate segment theorem. The other three students were assessed as successful in proof modification because they found other local counterexamples and successfully adjusted their proofs to cope with these local counterexamples.

Discussion

We have relied and built upon previous research on proof validation (e.g. Alcock and Weber 2005; McCrone and Martin 2004; Selden and Selden 2003) by addressing proof validation and modification involving local counterexamples at the secondary school level. Although several students did not succeed in proof validation/modification in our task-based questionnaire, these results are understandable if we take the difficulty of the task into consideration. We conducted our study in a lower secondary school, the participants already being familiar with the inscribed quadrilateral theorem that was necessary for proof modification. However, the Japanese national curriculum specifies that this theorem should be learnt in the tenth grade in upper secondary schools. Furthermore, if a simpler proof problem had been used in the questionnaire, such as the task in the first lesson, more students would likely have succeeded in proof validation and modification.

The students' failure in proof validation resulted mainly from their difficulty with drawing diagrams that satisfied the condition of the proof problem. We conjecture that this is a phenomenon that goes beyond our study. Our conjecture is based on the prevalent practice in schools, where diagrams of proof problems are usually provided by teachers and textbooks (Herbst and Arbor 2004; Herbst and Brach 2006). The students therefore have limited opportunity to produce diagrams by themselves through considering problem conditions, which consequently hinders their skill in proof validation. While this explanation is plausible, it has not been confirmed, so future research may need to address it.

As mentioned in the beginning of this chapter, proof-related activity consists of several phases, and we have focused on the particular phase that occurs after proof construction. In particular, this study defines proof validation and modification in terms of the discovery of local counterexamples by producing diagrams. This type of mathematical activity is important in school mathematics. For instance, de Villiers (2010) examines mathematical experimentation, which includes diagrammatic evaluation of conjectures and proofs, and lists its functions in the discipline of mathematics. He argues that "we need to explore authentic, exciting and meaningful ways of incorporating experimentation and proof in mathematics education, in order to provide students with a deeper, more holistic insight into the nature of our subject" (de Villiers 2010, p. 220). Therefore, the proof validation and modification activity examined in this chapter is significant because it allows students to experience an integration of experimentation and proof, through which they may build a richer image of mathematics.

To achieve this, we used a specific kind of task, proof problems with diagrams. This leads to a direct implication for teaching: teachers in secondary schools may use this kind of task to introduce proof validationand modification into their geometry classes. Although existing studies have suggested that undergraduates' capability in proof comprehension and validation could be improved by relevant training (e.g. Alcock and Weber 2005; Hodds et al. 2014), they have not explored the kind of task that could be used in the training. This study may offer an approach to address this issue because, as has been described in the chapter, proof problems with diagrams can be employed to introduce proof validation involving the discovery of local counterexamples. However, when using such tasks, teachers need to consider the potential difficulties described in this chapter and prepare appropriate lessons that help students produce diagrams that satisfy the conditions of problems.

There are several limitations in this study. The intervention reported on in this chapter was only implemented in a single class at a secondary school in Japan; it is thus important to reiterate the cycle of the implementation, analysis, and improvement of classroom-based interventions to develop more robust task sequences for facilitating the improvement of students' proof validation and modification skills. Finally, because the results of the task-based questionnaire are based on a single proof problem with a relatively small number of participants, future studies should use different tasks and conduct larger-scale surveys to further examine these results.

Acknowledgements This study is supported by the Japan Society for the Promotion of Science (Nos. 15H05402, 16H02068, and 26282039).

References

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

Common Core State Standards Initiative. (2010). *Common core state standards for mathematics*. Retrieved August 27, 2015, from http://www.corestandards.org/wp-content/uploads/Math_Standards.pdf.

Department for Education. (2014). *Mathematics programmes of study: Key stage 4 (National curriculum in England)*. Retrieved May 17, 2016, from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/331882/KS4_maths_PoS_FINAL_170714.pdf.

de Villiers, M. (2010). Experimentation and proof in mathematics. In G. Hanna, H. N. Jahnke, & H. Pulte (Eds.), *Explanation and proof in mathematics: Philosophical and educational perspectives* (pp. 205–221). New York: Springer.

Herbst, P., & Arbor, A. (2004). Interactions with diagrams and the making of reasoned conjectures in Geometry. *ZDM The International Journal on Mathematics Education, 36*(5), 129–139.

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

Hodds, M., Alcock, L., & Inglis, M. (2014). Self-explanation training improves proof comprehension. *Journal for Research in Mathematics Education, 45*(1), 62–101.

Inglis, M., & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education, 43*(4), 358–390.

Knuth, E. J. (2002). Secondary school mathematics teachers' conceptions of proof. *Journal for Research in Mathematics Education, 33*(5), 379–405.

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

Komatsu, K. (2017). Fostering empirical examination after proof construction in secondary school geometry. *Educational Studies in Mathematics, 96*(2), 129–144.

Komatsu, K., Tsujiyama, Y., Sakamaki, A., & Koike, N. (2014). Proof problems with diagrams: An opportunity for experiencing proofs and refutations. *For the Learning of Mathematics, 34*(1), 36–42.

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

McCrone, S. M. S., & Martin, T. S. (2004). Assessing high school students' understanding of geometric proof. *Canadian Journal of Science, Mathematics and Technology Education, 4*(2), 223–242.

Miyazaki, M., Nagata, J., Chino, K., Fujita, T., Ichikawa, D., Shimizu, S., et al. (2016). Developing a curriculum for explorative proving in lower secondary school geometry. In *Proceedings of the 13th International Congress on Mathematical Education*. Hamburg, Germany.

Reiss, K., Klieme, E., & Heinze, A. (2001). Prerequisites for the understanding of proofs in the geometry classroom. In M. van den Heuvel-Panhuizen (Ed.), *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 97–104). Utrecht, Netherlands.

Segal, J. (1999). Learning about mathematical proof: Conviction and validity. *Journal of Mathematical Behavior, 18*(2), 191–210.

Selden, A., & Selden, J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education, 34*(1), 4–36.

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

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

Stylianides, A. J. (2007). Introducing young children to the role of assumptions in proving. *Mathematical Thinking and Learning, 9*(4), 361–385.

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

Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education, 39*(4), 431–459.

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

Chapter 10
Classroom-Based Issues Related to Proofs and Proving

Ruhama Even

Abstract This commentary chapter focuses on two main issues related to the challenge of classroom teaching and learning of proofs and proving: (1) classroom-based interventions, and (2) teacher interventions in students' argumentation. The chapter uses these two issues to comment on the four chapters included in Theme 2 of the monograph, and concludes by suggesting another classroom-based issue that could profit from more systematic work.

Keywords Proof · Proving · Classroom · Interventions · Argumentation

Theme 2 of this monograph focuses on classroom-based issues related to proofs and proving: essential components of doing and learning mathematics. The four chapters included in this section deal with two main issues related to the challenge of classroom teaching and learning of proofs and proving: (1) classroom-based interventions, and (2) teacher interventions in students' argumentation. Below I use these two issues to structure my comments on these four chapters. I conclude my commentary by suggesting another classroom-based issue that could profit from more systematic work.

Classroom-Based Interventions

As pointed out in the four contributions to Theme 2, accumulating research suggests that many students encounter difficulties with proofs and proving. A typical response to research findings that reveal students' difficulties in mathematics is the design of interventions that address these difficulties and have the potential to offer better learning experiences to students. Two of the contributions in this section (Komatsu et al., this volume; Stylianides and Stylianides, this volume) belong to

R. Even (✉)
Weizmann Institute of Science, Rehovot, Israel
e-mail: ruhama.even@weizmann.ac.il

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_10

this line of scholarly work. They focus on the issue of classroom-based interventions whose aim is to address difficulties students face so that teaching of proofs and proving is improved. Each of the contributions uses a different approach to address this issue.

Stylianides and Stylianides approach the issue of classroom-based interventions by attending to the scarcity in the mathematics education literature of theoretical frameworks that can be used to design classroom-based interventions that aim to address problems of students' learning. The researchers address this lack in the literature by proposing a theoretical framework for designing classroom-based interventions that aim to address key and persistent problems of students' learning in the area of proofs and proving (and in other areas as well). Stylianides and Stylianides discuss and exemplify three characteristics, which they argue are crucial for any classroom-based intervention: (1) the inclusion of an explanatory theoretical framework, (2) a narrow and well-defined scope that enables a short duration, and (3) the inclusion of a mechanism that triggers and supports conceptual change. The first characteristic they propose, which highlights the importance of "understanding of how the interventions "work" or "can work" to support student learning" (Stylianides and Stylianides, this volume, p. 104) can be easily agreed upon. However, the other two characteristics would probably not receive a full consensus.

I will start with the third characteristic proposed by Stylianides and Stylianides. This characteristic, which refers to conceptual change, is strongly associated with a cognitive perspective on learning, centering on the acquisition of knowledge by individual students. However, it does not fit well with some of the other theoretical perspectives on learning that are commonly used in mathematics education (e.g., Sriraman and English 2009). For example, what would be an equivalent third characteristic in the case of a situated perspective on learning (Lave and Wenger 1991) where the focus is on changes in students' participation in the classroom's collective activities and practices? In cases where the focus is on changes in the classroom discourse (e.g., Sfard 2007)? It would be useful to examine what variations of Stylianides and Stylianides' "third characteristic" might be employed for different theoretical perspectives on learning commonly used by the community of mathematics educators.

The second characteristic in Stylianides and Stylianides' theoretical framework, which emphasizes the importance of a short duration for classroom-based interventions, is particularly thought-provoking, even when only a cognitive perspective on learning is adopted. The vast literature on misconceptions suggests that "[t]he reason why misconceptions are stubborn is that they are viable, useful, workable, or functional in other domains or contexts" (Fujii 2014, p. 454). Thus, one would typically expect an intervention that address a *persistent* problem of students' learning in mathematics, "such as common and hard-to-remediate misconceptions" (Stylianides and Stylianides, this volume, p. 108) to extend over a considerable period of time, especially when an explanatory theoretical framework is expected regarding how the intervention "works". More research is therefore needed to better understand which key and persistent problems of students' learning in the area of proofs and proving can be addressed by classroom-based interventions that satisfy the second characteristic, and what such a requirement might entail. The advantages of short interventions, as indicated in Stylianides and Stylianides' contribution, serve as a good incentive to devote future research efforts in this direction.

The classroom-based intervention described in Komatsu et al. (this volume) is in line with the second characteristic in Stylianides and Stylianides' theoretical framework. Komatsu et al. report on a short (two-lesson) intervention whose scope is narrow and well-defined, aimed at introducing proof validation and modification into secondary school geometry. In contrast with other studies, Komatsu et al.'s novel intervention deals with "not-completely-invalid proofs, namely proofs that are valid in certain domains, and invites students to determine whether there are cases to which the proofs are not applicable" (p. 134). In this study, the researchers used a specific kind of geometrical task: proof problems with diagrams that involve the discovery of local counterexamples. Komatsu et al. found that students' failure in proof validation was closely connected to their difficultly in producing diagrams that satisfy the conditions of a given problem. This result implies the need for designing an intervention that addresses that difficulty. Thus, more design work is required in order to produce a classroom-based intervention aimed at addressing students' problems in proof validation and modification that "work". Hence, in a way, one needs to design a chain of interconnected interventions, which is in itself an extended intervention. Time will tell if such classroom-based interventions would satisfy all three characteristics required by Stylianides and Stylianides' theoretical framework.

Teacher Interventions in Students' Argumentation

Classroom-based interventions, like any classroom teaching, involve teacher-student interactions. This is the focus of the other two contributions in this section (Goizueta and Mariotti, this volume; Pedemonte, this volume). Both contributions deal with teacher-student interactions with regard to proofs and proving, focusing on teacher interventions in students' argumentation.

Proving is sometimes associated with providing support for or refuting a mathematical conjecture that has been stated by someone else (e.g., Wiles' proof of Fermat's Last Theorem, a student's proof to the claim stated in the geometry problem in Komatsu et al.'s chapter). However, proving is often associated with other essential characteristics of doing mathematics and argumentation—activities such as exploring and generating conjectures and mathematical claims, confronting and evaluating alternative positions, and participating in discussions where mathematical arguments are constructed and critiqued.

The setting in Goizueta and Mariotti's study had the potential to include such characteristics. Students worked in small groups solving a problem; they were expected to explore a probabilistic situation and generate conjectures about a suitable mathematical model based on interpreting the elements and the relationships defined by the problem situation. They had to confront and evaluate alternative suggestions proposed by their group mates, and participate in discussions where the validity of a proposed mathematical model is constructed and critiqued using mathematical arguments connected to the given problem situation.

However, things did not always work as intended. Goizueta and Mariotti noticed that sometimes during small group work, students interpreted the teacher's interventions as signals to alter their conjecture regarding a suitable mathematical model. Here the validity of a proposed mathematical model was determined by the authority of the teacher rather than by analyzing the relationship between the model and the given problem situation. Thus, unintentionally, the teacher's interventions distracted students from analyzing the elements and the relationships defined by the problem situation in order to construct a mathematical model. Instead, only after the teacher approved a proposed model did students attend to the given problem situation and interpret it in accordance with the model validated by the teacher. Goizueta and Mariotti term this intriguing phenomenon *ex post facto* modeling.

Pedemonte also examined teacher-students' interactions, focusing on teacher interventions when students are engaged in constructing a proof. Similar to Goizueta and Mariotti, Pedemonte shows how student arguments might be affected by a teacher's interventions. However, in contrast with Goizueta and Mariotti's approach, Pedemonte examined not only cases when the teacher's interventions were not effective—or even counterproductive, as demonstrated in Goizueta and Mariotti's study—but also cases when the teacher's interventions were effective. By using Toulmin's model for analyzing teachers' interventions in students' argumentation, Pedemonte conjectured that to be effective, the teacher's interventions should not "interrupt the *cognitive continuity* between student's argumentation and proof. The teacher's intervention acts as a rebuttal in the student's argument invalidating his conception but not replacing it" (p. 127). Pedemonte adds that a teacher's intervention is ineffective if it "replaces the student's argument by her own argument" (p. 128).

Pedemonte's hypothesis regarding what makes teachers' interventions in students' argumentation effective and what makes them ineffective implies that before intervening, teachers need to understand the student's argumentation. This is commonly accepted by the community of mathematics educators as important, yet it is not an easy task because teachers' interpretations of students' work involves ambiguity and difficulty. Teachers "hear students through" their personal and social resources, such as the teacher's lesson plan, her knowledge about the nature and possible sources of students' misconceptions, her expectation from a specific student, and her own way of solving the mathematics problems she presented to her students (Even 2005). Thus, understanding what students are saying and doing should not be regarded as unproblematic or as something certain. This implies that providing effective teacher's interventions (in Pedemonte's terms) might be more difficult to achieve than perhaps is anticipated.

Conclusion

The four contributions in the Theme 2 section of the monograph deal with important classroom-based issues related to proofs and proving. Readers interested in the teaching and learning of proofs and proving will find in these contributions valuable information, insights, and ideas. Obviously, four contributions cannot cover all of the important classroom-based issues related to proofs and proving. Next, I will discuss another important issue that could benefit from more systematic research.

Classroom-based interventions—addressed by Stylianides and Stylianides (this volume) and Komatsu et al. (this volume)—are often associated with revising the mathematical content to be studied, and developing new curriculum materials. The important role of curriculum materials in shaping students' opportunities to learn mathematics is supported by recent research, which suggests that in many countries curriculum materials, and especially textbooks, greatly shape mathematics classroom instruction. Mathematics textbooks are often the main source that teachers use to plan lessons, choose the content to be taught, and the activities to be conducted (Eisenmann and Even 2011; Gueudet et al. 2013).

However, curriculum materials are only one factor involved in shaping students' opportunities to learn mathematics. Accumulating research suggests that students' opportunities to learn mathematics vary across classrooms taught by different teachers that use the same textbook (e.g., Ayalon and Even 2016; Even and Kvatinsky 2010). These studies highlight the prominent and indispensable role that teachers play in shaping how the curriculum is enacted in the classroom, and it underscores teachers' central role in determining the nature of the learning experiences provided to students—a role that no curriculum materials by themselves can fulfill. This includes prompting students to establish claims and justifications, encouraging them to critically consider different arguments, and modeling to students what constitutes acceptable mathematical proofs. The central role of the teacher in shaping students' opportunities to learn proofs and proving was vividly demonstrated in Goizueta and Mariotti (this volume) and in Pedemonte (this volume).

Moreover, research also shows that differences in students' opportunities to learn mathematics, including proofs and how to prove, occur not only between classrooms of different teachers, but also between classrooms taught by the same teacher and that use the same textbook (Ayalon and Even 2016; Eisenmann and Even 2011). These studies clearly underscore the key role that classrooms—together with the teacher and the curriculum materials—play in shaping students' opportunities to learn mathematics in general and proofs and proving in particular.

For example, Ayalon and Even's (2016) study comprised two case studies in which each of two teachers taught in two 7th grade classrooms, and all four classrooms used the same textbook. Both teachers followed rather closely the teaching sequence suggested in the textbook, and the classwork in all classrooms consisted almost entirely of work on tasks from the textbook. However, the two teachers adopted different teaching approaches in relation to engaging in proofs and proving. One teacher prompted students to hypothesize, pressed for justifications, as well as encouraged critical listening to others' claims, seldom modeling acceptable ways of proving. In contrast, the other teacher modeled adequate ways of building complete proofs, and put little emphasis on involving students in this activity.

Additionally, each teacher's two classrooms had different characteristics. In the case of the former teacher, one classroom was cooperative, with highly motivated students; the other classroom was active but sometimes the students had trouble engaging in mathematics. In the case of the latter teacher, one classroom was characterized by active student participation, whereas the other classroom was characterized by lack of student participation and frequent disciplinary problems.

Analysis revealed that students' opportunities to engage in proofs and proving were rather similar in both classrooms of the latter teacher, who adopted a teacher-centered "modeling" approach. A similarity was also found regarding students' opportunities to engage in proofs and proving in both classrooms of the former teacher—who avoided intervening in students' argumentation and refrained from modeling proving—when the class dealt with a topic that mostly required use of inductive reasoning (investigating algebraic expressions). However, substantial differences were found between the two classrooms of this teacher with regard to students' opportunities to engage in proofs and proving, when dealing with a topic that required extensive use of deductive reasoning (equivalence of algebraic expressions). An analysis revealed that it was the interplay among a teacher's fundamental teaching approach, the specific characteristics of each of her two classrooms, and the contrasting characteristics of the two topics that greatly contributed to the similarities and differences found.

Most research studies that examine students' opportunities to learn mathematics usually focus on one factor (typically the curriculum or the teacher), contributing important knowledge about its role in shaping students' opportunities to learn mathematics, as was demonstrated by the Theme 2 chapters. However, not much is known about how students' opportunities to engage in the mathematics classroom in mathematics in general, and in proofs and proving in particular, are shaped by the interplay of several factors, which are intrinsic to classroom teaching and learning. To date, the interplay of such factors has received little research attention and researchers often consider them as "noise". Thus, more research is needed to advance our theoretical knowledge about how the interplay among the characteristics of the curriculum, the teacher, and the classroom shape students' opportunities to engage in proofs and proving in the mathematics classroom.

References

Ayalon, M., & Even, R. (2016). Factors shaping students' opportunities to engage in argumentative activity. *International Journal of Science and Mathematics Education, 14,* 575–601.

Eisenmann, T., & Even, R. (2011). Enacted types of algebraic activity in different classes taught by the same teacher. *International Journal of Science and Mathematics Education, 9,* 867–891.

Even, R. (2005). Using assessment to inform instructional decisions: How hard can it be? *Mathematics Education Research Journal, 17*(3), 45–61.

Even, R., & Kvatinsky, T. (2010). What mathematics do teachers with contrasting teaching approaches address in probability lessons? *Educational Studies in Mathematics, 74,* 207–222.

Fujii, T. (2014). Misconceptions and alternative conceptions in mathematics education. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 453–455). Dordrecht, Heidelberg, New York, London: Springer.

Gueudet, G., Pepin, B., & Trouche, L. (2013). Textbook design and digital resources. In C. Margolinas (Ed.), *Task Design in Mathematics Education: Proceedings of ICMI Study 22* (pp. 325–336). Oxford, UK: ICMI.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.

Sfard, A. (2007). When the rules of discourse change, but nobody tells you: Making sense of mathematics learning from a commognitive standpoint. *The Journal of the Learning Sciences, 16*(4), 565–613.

Sriraman, B., & English, L. (Eds.). (2009). *Theories of mathematics education: Seeking new frontiers*. Berlin: Springer.

Part III
Cognitive and Curricular Issues Related to Proof and Proving

Chapter 11
Mathematical Argumentation in Pupils' Written Dialogues

Gjert-Anders Askevold and Silke Lekaus

Abstract In this chapter, we present some results from a project about mathematical argumentation and proving in the form of dialogues. Tasks were prepared in the form of written dialogues between two imaginary pupils discussing a mathematical problem, and pupils were invited to write their own dialogues continuing the mathematical discussion. An analysis of dialogues about fractions written by 33 pupils from two classrooms in Norway in Grades 5 and 6 (10–12-year-olds), working in small groups, revealed that many of the 5th grade pupils used forms of argumentation supported by visual representations of fractions, while the 6th graders used more rule-bound approaches based on conversion. The analysis showed that three of ten groups in 6th grade used both diagrammatic and narrative argumentation in contrast to 5th grade where half of the groups were able to use these two kinds of argumentation. Those groups who made use of both types of argumentation were most successful in their argumentation. We relate these findings to the theory of relational and instrumental understanding in mathematics.

Keywords Mathematical argumentation · Proof · Imaginary dialogues

Introduction

There is wide evidence of the fact that many pupils and even students at college level have misconceptions about the nature of mathematical proof and cannot actively construct their own proofs. Stylianides et al. (2017) gave a detailed overview of research studies in this area giving such evidence. Many students apparently fail to understand that a proof ensures the truth of a statement and makes further examination unnecessary. It seems to be difficult for students to move from empirical thinking, as done in everyday life, to logical-deductive thinking as required in understanding and producing mathematical proof: "To construct a proof

G.-A. Askevold · S. Lekaus (✉)
Western Norway University of Applied Sciences, Bergen, Norway
e-mail: slek@hvl.no

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_11

requires an essential shift in the learner's epistemological position: passing from a practical position (ruled by a kind of logic of practice) to a theoretical position (ruled by the intrinsic specificity of a theory)" (Balacheff 2010, p. 118). Healy and Hoyles (2000) studied the proof conceptions of high-attaining 14–15 year old pupils. They found that many pupils produced empirical arguments instead of proof. An interview study with the same pupils revealed that many of them were aware of the fact that these were not proofs, but also that the pupils were not able to produce more formal arguments. Work by Knuth et al. (2009) with pupils in grades 6–8 revealed similar results and suggests the need for more research about the relationship between pupils' proof production competencies and their proof comprehension competencies.

The role of proof and proving in school mathematics has received increased attention in both the mathematics education community and mathematics education reform initiatives (Knuth et al. 2009). The role that proof can play in providing explanations and understanding of mathematical content is discussed by Hanna (this volume). As stated by Harel and Sowder (2007), it seems that "at least some of the deficiencies in students' acquisition of more sophisticated proof schemes may stem from the lack of opportunity to engage in proof-fostering activities" (p. 828). In Norway, where this study was conducted, Mellin-Olsen (1996) noted the strong focus on computational exercises and rote learning in textbooks and classroom practice. In recent years, we observed changes in the Norwegian national curriculum intending to make pupils active users of mathematics. Curriculum development has resulted in emphasis on five basic skills (oral skills, reading, writing, digital skills and numeracy) that are to be developed in connection to all mathematical subjects and which can be related to proving activities, though the word proof is not mentioned in the Norwegian curriculum for grades 1–10 (pupils aged 6–16). Proof in mathematics is not only a (potential) curricular topic, but can serve a number of functions like explanation, systematization, discovery and communication, as identified by de Villiers (1990). A similar list of categories of intellectual needs in the learning and teaching of mathematics is given by Harel (this volume). These functions of proof and proving pervade the descriptions of all basic skills in the Norwegian curriculum and are thus intended to be a part of all mathematics teaching and learning. In addition to "asking questions", "communicating ideas" and "describing and explaining a process of thought and putting words to discoveries and ideas", the curriculum requires the pupils to learn to build up "argumentation with help from informal language, precise terminology and the use of concepts" (Ministry of Education and Research 2013).

In order to contribute to the implementation of the above-mentioned curricular intentions in the classroom, we started to explore the potential of pupils' dialogue writing on different topics in mathematics. Gholamazad (2007, p. 266) states: "Writing down the dialogue may provide students an opportunity to reflect on their thinking process and to organize it in a convincing way. In this perspective, the dialogue can be considered as an intermediate stage between having an overview of a proof and writing a formal mathematical proof." In this chapter, we present results from an analysis of pupils' written dialogues about fraction problems. Tasks were prepared in the form of written dialogues between two imaginary pupils discussing a mathematical problem, and pupils were invited to write their own dialogues continuing the mathematical discussion. This was a modification of the method of imaginary dialogues used by Wille (2011). In addition to exploring the potential of dialogue writing to support pupils' ability to build a mathematical argumentation, we analyzed pupils' construction of arguments and their conceptions about proofs as expressed in their written texts.

Theoretical Framework

Hemmi (2008) applied the social practice theories by Lave and Wenger (1991) and Wenger (1998) to students' learning processes of mathematical proof and proving. Proofs were seen as artefacts in mathematical practice that are mostly learned by enculturation. This refers to the process of newcomers becoming participants in a community of practice by gradually learning how to talk, act and use artefacts. A key notion in this context is the condition of *transparency*. Hemmi found that the teaching and learning of proofs as artefacts in mathematical practice require a balance between their visibility (by focusing on their logical structure, function etc.) and their invisibility (by focusing on explaining, justifying and convincing without explicitly connecting these to the construction of proofs). Hemmi called this an "unproblematic" use of proof (p. 414) in contrast to an explicit focus on proof. This was based on work by Adler (1999) who studied talk as an artefact in mathematical practice. She revealed a teaching dilemma of transparency concerning the use of mathematical language. According to her study, explicit teaching of mathematical language, although necessary, can obstruct the flow in a mathematical discussion when pupils understand the mathematics that they are discussing, but are expressing it incorrectly.

A related approach to students' learning of proof can be found in the work of Blum and Kirsch (1991) who discussed the significance of *preformal proof*, which they defined to be a "chain of correct, but not formally represented conclusions which refer to valid, non-formal premises" (p. 187). The construction of preformal proof is a way of arguing where the formal aspects of proof are not paid attention to, i.e., remain invisible. This can make it an "unproblematic" use of proof (Hemmi 2008). The pupils' own writing of dialogues containing mathematical argumentation as done in the project described in this chapter can be considered a first step for young learners to be guided to the construction of preformal arguments and proofs, the concept of proof and its formal aspects not being visible.

In our approach to imaginary dialogues, the pupils were always encouraged to support their writing process by using sketches and drawings. Hanna (this volume) states that visual representations may enhance explanatory proofs and understanding in mathematics. As one of the necessary preconditions for students to understand, judge or even construct preformal proofs, Blum and Kirsch (1991) identified the need "to place value on manifold kinds of representations of mathematical content, especially to stress reality-oriented basic ideas and to impart geometric intuitive basic conceptions" (p. 200). The role of visual representations for the development of mathematical argumentation was thoroughly discussed by Dreyfus et al. (2012). They referred to several studies that found that visual representations play an important role even in the work of professional mathematicians, even though they are not precise and only represent special cases. Among other aspects, they emphasized the need for a "fluent interplay between analytical rigour and (often visually based) intuitive insight" (p. 194). The important role of drawings for the development of pupils' mathematical understanding in proving processes is also discussed by Komatsu et al. (this volume).

For the analysis of the argumentation found in the pupils' written dialogues, the work of Krummheuer (1999, 2013) provided a powerful tool. As described in Krummheuer (2013), visual representations in the form of inscriptions play an important role for mathematical argumentation. He investigated the relationship between diagrammatic and narrative argumentation in the development of mathematical thinking of children of kindergarten and early primary school age. Based on the work of the linguist Peirce (1978) about inscriptions and diagrams, Dörfler (2006) and Krummheuer (2013) developed the notion of *diagrammatic argumentation* in mathematics. Diagrammatic argumentation consists in the production, use and transformation of diagrams. Such diagrams are inscriptions together with a "(conventional) system of rules concerning their production, use and transformations" (Dörfler 2006, p. 202). Another form of argumentation, which was identified by Krummheuer (2013), is *narrative argumentation* which is characterized by "the invariability of a sequence of sentences" (p. 251). As in a story, the order of events cannot be changed. The presentation of a narrative argumentation "proceeds mainly by verbalization" and there are little or no visualizations (Krummheuer 1999, p. 338). The mathematical concepts underlying such argumentation are often not introduced explicitly. The ability to use diagrammatic argumentation seems to precede the development of narrative argumentation. Krummheuer (2013) noticed that narrative and diagrammatic argumentation appeared blended in the argumentation of young school children but that such a combined appearance of the two forms of argumentation seemed "to have only little stability as an intermediate stage in the negotiation process" (p. 260). Working with a problem, pupils argued both by transforming a diagram (diagrammatic argumentation) and by discussing a problem verbally (narrative argumentation), but they rarely combined these two kinds of argumentation other than just briefly, as a short intermediate stage in the process of finding an answer.

Krummheuer (2013) further conjectured that a combination of diagrammatic and narrative argumentation that is restricted to describing "the set of rules that prescribes the production and transformation of a mathematical diagram" and providing a "sequence of sentences that narratively describe only the concrete application of these rules" could lead to algorithmic, mechanical ways of thinking (p. 260). This means that the two forms of argumentation would not support each other in the building of a mathematical argumentation, but that the narrative elements would only be used to describe the rules for how to create and transform diagrams. That type of understanding, restricted to the knowledge of how to apply rules in a restricted range of known situations, was called *instrumental understanding* by Mellin-Olsen (1984), in contrast to *relational understanding* that, roughly speaking, includes knowledge about why the rules work and how they are connected to the underlying mathematical concepts and objects. Skemp (1976) gave a thorough description of these two forms of mathematical understanding.

In their study of pupils' proof conceptions, Healy and Hoyles (2000) used the term narrative argumentation in a different meaning than Krummheuer, referring to (often pupils' own) argumentation in everyday language, possibly supported by drawings, and avoiding the use of algebraic language. Healy and Hoyles found that many pupils held simultaneously two conceptions about proofs. The type of argument that they expected the teacher would appreciate contained formal, algebraic language and differed from the type of argument that they would use when trying to convince themselves. They found that pupils preferred narrative argumentation due to its greater explanatory power and that pupils were more successful in building deductive argumentation, when using narrative argumentation. This is related to Hemmi's notion of transparency: the narrative style of argumentation without the obligation to use formal mathematical language can be a form of unproblematic use of proof.

By the pupils' writing of dialogues as described in this project, we wished to encourage their mathematical argumentation in a non-formal environment, explicitly allowing the use of non-formal language typical for oral conversations. This kind of task can be regarded a first step for young learners towards the construction of preformal arguments and proofs, the concept of proof not being visible.

Method

This chapter is based on an analysis of mathematical texts written by pupils from two classrooms in grade 5 and 6, i.e., pupils aged 10–12, in Norway. We presented the plans for the project to a group of dedicated mathematics teachers underlining the importance of an investigative and collaborating atmosphere in the mathematics classroom. Two volunteering teachers opened their classrooms for our project. The tasks given to the pupils had the form of short written dialogues of two imaginary pupils having a conversation about a mathematical problem. In addition, we as researchers presented the tasks to the pupils gathered in a whole-classroom situation. The pupils in the classroom were asked to continue writing the dialogues in small groups while they were investigating the mathematical problems. The pupils worked on the tasks with both their regular teacher and us being present as assistants, answering general questions about the task, but not giving mathematical guidance towards a solution. Based on research about the significance of visual representations (Dreyfus et al. 2012), the pupils were also encouraged to use drawings to support their thinking. This way of working with mathematical problems was inspired by the method of imaginary dialogues (Wille 2011, 2013). In contrast to Wille's study, in which the pupils had to work individually, pupils in our study worked in small groups of 2–4 pupils (with the exception of one pupil who asked to write individually). We arranged the classes in this way because it was a new type of assignment and we wished the writing of dialogues to become a good experience for the pupils where they could rely on each other and have many ideas to build upon. A total of 33 pupils participated in the two classrooms. We collected 16 dialogues, six in 5th grade and ten in 6th grade. We only had resources to videotape two groups in each classroom. In this chapter, we mainly report on our analysis of the dialogues written by the pupils who were giving the most convincing arguments; unfortunately these pupils were not videotaped.

The Given Task

The task analyzed here was the pupils' second encounter with mathematical reasoning in the form of dialogues. In the previous week, the pupils had worked with a task designed by Wille (2011) about the limit of the geometrical series (Lekaus and Askevold 2015). This time we expected them to be more focused on the mathematical argumentation since the task design was familiar. The new introductory dialogue started with a short conversation between two fictive pupils, called Petter and Ragnhild, about what is more desirable to get, 1/10 or 1/3 of a cake. The dialogue then turned to a more formal problem of ordering some given fractions from least to greatest:

Petter: Yes, with fractions it is like this, they are smaller, when the number in the denominator is bigger. But listen, Ragnhild, this is only true when the number in the numerator is 1. What if we have different numbers in the numerator? Does it make any difference?

Ragnhild: Why do you ask so difficult questions? But perhaps you are right, let us investigate a little. I want to check which of the following fractions is the biggest and the smallest one $\frac{3}{4}$ $\frac{2}{3}$ $\frac{9}{12}$

Petter: Let us also look at the following five fractions: $\frac{3}{7}$ $\frac{4}{5}$ $\frac{4}{8}$ $\frac{9}{7}$ $\frac{3}{5}$

Ragnhild: Fine, we have eight fractions. Let us order them from least to greatest.

The fractions were chosen in a way that made it arithmetically difficult for the pupils to find a common denominator for all the given fractions. By this, we wanted to test which approaches the pupils would turn to when a purely rule-bound approach to the solution was not easy to apply.

The pupils in the two classrooms had some experience in adding and converting fractions and knew how to compare fractions that have the same denominator. The pupils were also familiar with some visual representations of fractions used by their textbooks and teachers. The 5th and 6th graders had similar prior knowledge about fractions, but the 6th graders had some more computational practice.

Analysis

The pupils expressed their argumentation in different semiotic systems, like text, calculations and drawings. However, the dialogues and the corresponding drawings were produced on separate sheets of paper, so we could not always be sure which of the sketches the pupils were referring to in their dialogues, but we had in most cases strong evidence for the correspondence. The dialogues written by the pupils are fictive conversations, and the fictive characters Petter and Ragnhild might show less knowledge than the pupils writing the dialogue actually have. It seems unlikely though that the pupils let both protagonists demonstrate less knowledge than the pupils themselves had. In the analysis, we therefore treated the demonstrated knowledge of the fictive characters in the dialogues as a demonstration of the pupils' own knowledge and argumentation skills. We analyzed whether the argumentations had characteristics of narrative or diagrammatic argumentation and in which way the use of visual representations supported the development of mathematical argumentation. This included the variety of representations that were used and the pupils' ability to switch between and combine different representations or diagrams. In the cases in which the pupils did not come to a conclusion in the argumentation process, we examined whether this could be explained by a lack of ability to merge the two forms of argumentation or by an instrumental application of either one.

The data for this chapter consist of the six dialogues written by the 5th graders and the ten dialogues written by the 6th graders. In our analysis, we used both the term representation and the term diagram. Representations are inscriptions symbolizing mathematical objects, in our case fractions. We considered representations to be diagrams when transformations were visible in the drawings or other inscriptions or when the pupils clearly connected or applied a set of rules to a given representation. Consequently, we only speak about diagrammatic argumentation when the argumentation was based on the transformation of inscriptions. In cases when the written dialogues were accompanied by inscriptions (e.g., drawings or fractions) without any sign of transformation, we considered these to be narrative. Dialogues that had a narrative form were divided into two groups: those that contained narrative argumentation in the sense of Krummheuer (2013) and those that displayed a conversation between the protagonists without any mathematical argumentation.

Results

All groups of pupils explored the mathematical problem and arranged some of the fractions in the correct order. In particular, all groups managed to identify the biggest fraction. Our analysis showed that the 5th and 6th graders used different forms of argumentation. They built their argumentation on either arithmetic or

Table 11.1 A summary of our findings

	Use of conversion	Arguments based on visual representations	Use of both diagrammatic and narrative argumentation
6th graders	6 of 10	1 of 10	3 of 10
5th graders	3 of 6	6 of 6	3 of 6

visual representations of fractions. The 5th graders strongly depended on visual representations, while the 6th graders mostly used approaches based on conversion of fractions. The groups who used both diagrammatic and narrative argumentation were the most successful in building an argumentation. A summary of our results is found in Table 11.1.

Sixth Grade

All groups of pupils wrote a dialogue examining the task and found at least partial solutions ordering some of the fractions, but only in one of the dialogues from both classes the task was completed successfully. It was written by a boy from 6th grade, who we here call Ali, who asked permission to work individually. Ali chose to approach the problem arithmetically using conversion of fractions. He constructed a dialogue with a clear narrative structure, telling a story with an introduction, a main part and a summary. The first part of the dialogue was constructed as a negotiation between the fictive protagonists about finding a proper method for arranging the fractions in correct order. Early in the dialogue, Ali made the protagonist Petter observe that it is "not possible" to find a common denominator. It is not clear whether Ali was genuinely convinced that this was (theoretically) impossible or only impossible for him, in practice, to carry out. Ali chose instead the uncommon method of finding a common numerator, which is quite manageable for the given fractions. The crucial part of Ali's written dialogue is shown below, showing the shift in the mathematical process where the change of strategy takes place from finding a common denominator to a common numerator.

Petter	Okay, we have to get common denominator
Ragnhild	Yes sure!
Petter	But I don't believe that it is possible.
Ragnhild	What do you mean by that?
Petter	7, 5, 8, 7, 5, 4, 2 and 9 ~~7, 5, 8, 7 and 5~~ can't get a common denominator ((Ali's own crossing-out))
Ragnhild	Are you sure?
Petter	Yes, it's not possible!
Ragnhild	Hmm…
Petter	But maybe…
Ragnhild	What do you mean?
Petter	Maybe we can get a common numerator!

It was not clear whether Ali invented this method by himself or whether he learned it from a more competent person. The dialogue suggested certainty and comprehension regarding the procedure. It seemed that Ali was confident about the mathematics and the method. As observed by Wille and Boquet (2009) in similar cases, the fictive protagonists do not have equal roles: Petter was given the role of teaching Ragnhild. Nevertheless, Ali made the fictive character Petter praise the good collaboration several times. This might reflect classroom norms that ask pupils to encourage peers.

In the main part of the dialogue, the argumentation departed from the standard procedure, the use of a common denominator, but an explanation was given that made this departure plausible. This is a characteristic of narrative argumentation (Krummheuer 2013). The explanation of the details of the new procedure of using a common numerator had rather the form of diagrammatic argumentation; the inscription underlying the diagram was a numerical representation of equivalent fractions. All necessary transformations of diagrams needed to fulfil the argument were included in the text; see Fig. 11.1 for an example of these. The rationality of the procedure was explained by the character Ragnhild in a type of summary: "(…) the fraction with the least denominator (is) the biggest fraction, when the numerators are the same." The chosen path of argumentation, including both diagrammatic and narrative arguments, may reflect a relational understanding of fractions: Ali used standard conversion of fractions in an uncommon way explaining all the necessary steps in the argument. The resulting dialogue had the characteristics of a preformal proof: it had a logical structure and all steps were justified in a mathematically correct way, though the style and language were informal.

In only one of the dialogues written by the 6th graders the possibility of using visual representations was explicitly mentioned. This dialogue was written by three boys, who seem to have been experimenting with different strategies during the

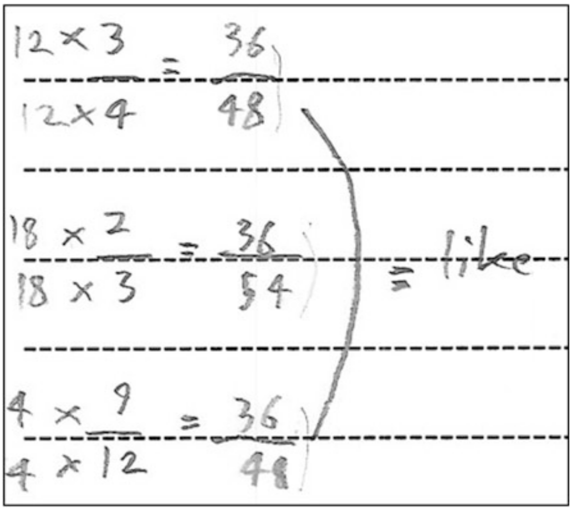


Fig. 11.1 "Like" means equal

writing process. They had the initial heuristic strategy of trying to reduce the problem to a simpler one, by identifying the biggest, the smallest and the second-smallest fraction. We sensed some uncertainty when they made their protagonists say that they "believe" that 3/7 is the smallest and 4/8 is the second-smallest fraction. Though there were only few inscriptions in this dialogue, we found a glimpse of diagrammatic argumentation in some notes that showed calculations of equivalent fractions. Some drawings of fraction circles accompanied the written dialogue, as shown in Fig. 11.2. They represented all the given fractions, but there was no evidence that these were transformed and used as a basis for diagrammatic argumentation. We regarded them therefore as visual representations, not diagrams.

The fact that the fraction circles were only mentioned briefly in the text, followed by the suggestion to find common denominators, indicates that the pupils did not manage to make use of them. The reason might be that the drawings were too inaccurate.

The following part of the dialogue had the characteristics of a narrative argumentation. It started with the fictive protagonists agreeing on the new strategy of finding a common denominator which was presented as if it was a well-known concept to the protagonists, and thus to the pupils themselves, that required no further justification. Explanations were rather vague ("48 cannot be common denominator because there is a 5") and required the addressee to have some knowledge about the concepts involved. It may have been an obstacle for solving the assignment that the pupils did not try or manage to combine the different approaches.

We take a short look at the obstacles other 6th graders met in their argumentation process: Some pupils developed short isolated paths of diagrammatic argumentation based on equivalent fractions, but they did not manage to relate different diagrams

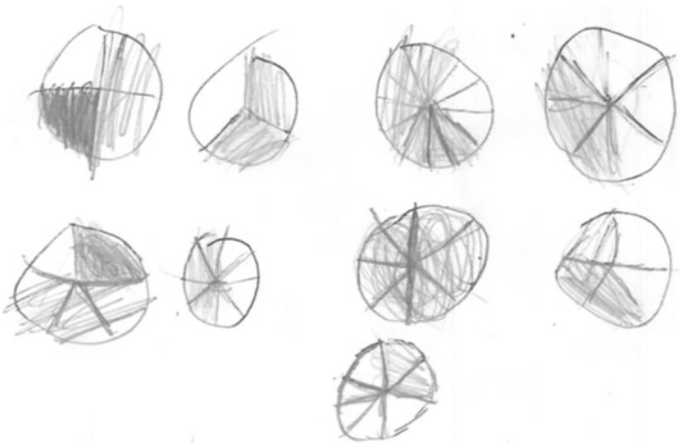


Fig. 11.2 Representation of the given fractions

to each other and to build their arguments on more than one diagram at a time. For example, they used conversion of fractions in order to compare separate groups of fractions that can have common denominator 12 or 35. This approach was found in five of the ten dialogues. This separation into two groups with a rather obvious common denominator made us believe that these pupils had rather an instrumental understanding of the concept of fractions, handling well the technical procedure of conversion, but not managing to build a complete argumentation on their partial results.

Other dialogues were narrative in style with a conversation between the protagonists developing around the task, but they presented no real mathematical argumentation. In these dialogues, we found opaque explanations like "I looked at the denominator and a little bit at the numerator, too", which were not backed up by the use of diagrams. We also found some pupils simply guessing the correct order of fractions, expressed by their protagonists saying "I believe that (…)". We also found traces of an authoritarian proof scheme (Harel and Sowder 2007) in some of the 6th graders' dialogues, when they let their protagonists appeal to an authority, here represented by a more competent peer, by saying "I trust you" or "you are the smarter one".

Fifth Grade

All the 5th graders based their argumentation strongly on visual representations. We take a closer look at two examples:

The first dialogue written by two girls consisted of three parts: an initial suggestion to use cake diagrams, an arithmetic approach involving conversion of fractions and a new visual approach, this time by equally long number lines divided into sections. In this dialogue, we found a mixture of narrative and diagrammatic argumentation. The initial considerations involving conversion of fractions were purely narrative and were not supported by any diagrams. The procedure was described only verbally and seemed to be fully agreed upon by both the fictive protagonists, stating no further grounds. Though the imprecise terminology "multiplication" was applied, the fractions 3/4 and 2/3 were converted correctly in order to compare them to 9/12 (see the excerpt of the girls' written dialogue below). We may sense some uncertainty about simplifying fractions, when 4/8 was reduced to 2/4, not 1/2, before comparing it to 3/5.

Ragnhild: Isn't it easier to convert the fractions?
Petter: Yes, because if we multiply 3/4 with 3 we get the same as 9/12, as long as we multiply the same above and below.
Ragnhild: Then we try with 2/3. 2/3 multiplied by 4 is 8/12 and that means that 2/3 is less than 9/12 and 3/4.

A closer analysis of the seemingly purely arithmetic part of the dialogues showed nevertheless that some visual representations were used also here to support the argumentation. The girls concluded that 3/5 is bigger than 2/4, for which we could not find any argumentation in the text, but only in their sketches of rectangles and number lines. Although the girls had knowledge about how to convert fractions, they did not combine conversion with the visual approach, and some inaccurate drawings made the girls draw incorrect conclusions, such as 3/7 being more than a half.

The major part of the written dialogue contained little argumentation and served rather as a place to summarize the results of diagrammatic argumentation based on transformations of rectangular cake diagrams and number lines divided into sections that were found on separate sheets of paper. The various approaches remained separate from each other in the written argumentation, but the results of the process were summarized in a joint number line, as shown in Fig. 11.3. From this figure and the last line of the dialogue ("I wonder which one is bigger of 4/8 and 3/7") we observed that it was not obvious to the girls that 4/8 is the same as one half or that 3/7 is smaller than one half. In this case, when inaccuracy of the diagrams made it impossible to draw a conclusion, the strong focus on visual representations seems to have been an obstacle in the argumentation process.

The second dialogue also written by two girls was accompanied by similar visual representations. These girls used pizzas, squares, number lines, and some arithmetic methods to visualize and compare fractions. Two paper sheets were covered in number lines divided into three, four, five or twelve parts. In three places we found the pupils using these in order to visually compare fractions pairwise. The drawings of pizzas and squares were few, and we did not find traces of the pupils manipulating them; we regarded them therefore as visual representations, not diagrams. The dialogue started by justifying that 9/7 is the biggest fraction since the numerator is two bigger than the denominator. Then the girls made the protagonists

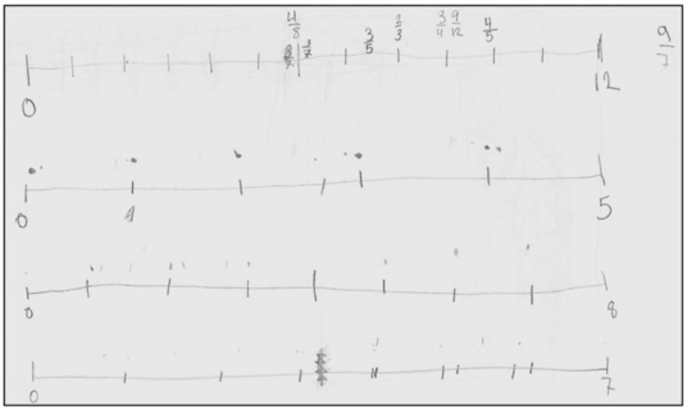


Fig. 11.3 Results of diagrammatic argumentation

apply a similar argumentation to compare fractions pairwise: they erroneously concluded that 3/5 is bigger than 9/12, using the incorrect argument that the numerator was only two smaller than the denominator, while the difference was three for the other fraction, not taking into account the size of the pieces. Nevertheless, the pupils recognized equivalent fractions, ignoring their own arithmetic rule involving differences. They possibly considered their rule again at the end when they suggested that the two fractions 4/5 and 3/4 might be equal. Based on visual representation, in this case number lines, they decided that 4/5 is bigger.

The appearance of this dialogue was narrative, but the argumentation had rather the characteristics of a series of short diagrammatic argumentations (pairwise comparison of fractions) that were supported by only one type of diagram: number lines divided into sections. We also found simplification of fractions once but the pupils did not explore this any further. The girls made an attempt to find and formulate a rule involving the number of missing pieces in the numerator based on experiments with the chosen diagram. It seems thus that the use of diagrams supported an investigative approach to the task though the resulting rule was not correct.

Though both dialogues were accompanied by drawings and one of them mentioned visual representation in the text, the longest parts of the dialogues were presented as a sequence of considerations that were seemingly developed using arithmetic methods. There was however strong evidence that visual representations were used while writing almost all parts of the dialogues and partly also were more convincing to the pupils. This might indicate that visual argumentation was the most manageable tool for the pupils, but that they might have thought that an arithmetic argument was the expected one or a more powerful tool. This seems to be consistent with findings by Healy and Hoyles (2000) that pupils preferred arguments expressed in everyday language combined with examples and diagrams due to their explanatory power, while they believed that the use of formal algebraic arguments was expected by their teachers.

Discussion

We found notable differences between the methods applied by the pupils from the two classrooms. The 5th graders used mostly arguments based on visual representations, and we found some variety in the chosen representations. They also made some use of conversion of fractions. The pupils seemed genuinely interested in investigating the problem; they did not express the need for an authority to support their method or to check their answer. The 6th graders produced few drawings and then only fraction circles. Only one group of pupils mentioned these explicitly in their written dialogue. Six of ten groups used arguments based on conversion of fractions. In various cases, the pupils referred rather vaguely to procedures or rules that they had learned or applied earlier. The fact that pupils referred to rules and procedures without further justification might suggest that the application of these procedures without justification was usually accepted by classroom norms. In some of the dialogues, the pupils mentioned mathematical notions related to fractions, but they did not put them in relation to each other and failed to build an argumentation on those: "we have to look at the denominators, and a little bit at the numerators, too". We consider this an illustration of what Mellin-Olsen (1984) called an instrumental understanding of fractions. Those pupils seemed to master only some of the computational procedures concerning fractions.

The reasons for the occurring differences in the pupils' argumentation in the two classrooms are not clear and we can only speculate about the reasons. As found by Healy and Hoyles (2000), pupils seem to note an increasing expectation to use more formal mathematical language when they progress in their school careers, and our findings might reflect this. The differences might also reflect diverse teaching styles applied in the two classrooms, with the 5th graders possibly being more used to mathematical investigation and the use of visual representations.

Only one pupil reached a complete and correct solution, using a creative method based on equivalent fractions. His solution contained sufficient justification and had the characteristics of preformal proof as defined by Blum and Kirsch (1991). All of the pupils' written dialogues had a narrative structure, i.e., a plot with a given order of actions carried out by the fictive protagonists, but not all of them had the characteristics of a narrative argumentation, as defined by Krummheuer (2013). In several dialogues, no real mathematical argumentation was presented. Instead, they contained vague references to rules concerning numerators and denominators. The pupils who supported their reasoning by diagrammatic argumentation in the sense of Krummheuer (2013), using numerical or visual diagrams, were more successful in developing and conveying their argumentation. Not all of the diagrammatic argumentation could be found mentioned in the written dialogues, but we could see in several cases that diagrammatic argumentation had taken place since a development of diagrams was visible in the pupils' drawings and drafts on separate sheets.

The most important obstacles in the argumentation process seemed to be an overemphasis on the visual representations and a lack of the ability to combine different pieces of diagrammatic argumentation to form a longer chain of argumentation. Those pupils who were strongly committed to the visual seemed to display a beginning relational understanding of the concept of fractions, but inaccuracy of drawings prevented them from getting the correct result when the fractions were of nearly similar size. In these cases, the development of argumentation was obstructed by a lack of interplay between analytical rigor and visually based intuitive insight as discussed by Dreyfus et al. (2012): the chosen diagrams gave imprecise information in the case of fractions of similar size, which created an obstacle to the argumentation when the pupils were not combining the visual with numerical methods.

When the pupils did not combine diagrams in their argumentation, the order of their investigation seemed arbitrary. The order of the investigated diagrams could thus be changed, though not the verbal sentences of the dialogue, i.e., the order of events in the story about the protagonists that was built around the mathematical argumentation. In these cases, the use of several diagrams provided no development in the argumentation and several strands of argumentation remained standing side by side. We see a connection between this and the conjecture by Krummheuer (2013) that by a lack of combination of the two types of argumentation their potential is not fully utilized.

Despite these difficulties, we found that the writing of dialogues between fictive protagonists helped the pupils to structure the investigation of different representations and diagrams and to formulate pieces of mathematical argumentation, one pupil even formulating a preformal proof. Other aspects of dialogue writing that could be examined further are how pupils use their language to formulate their arguments and which roles they make the protagonists take in the argumentation process. We expect these issues to give insight into to pupils' beliefs about mathematics and the knowledge and skills needed for engaging in mathematical argumentation.

References

Adler, J. (1999). The dilemma of transparency: Seeing and seeing through talk in the mathematics classroom. *Journal for Research in Mathematics Education, 30*(1), 47–64.

Balacheff, N. (2010). Bridging knowing and proving in mathematics: A didactical perspective. In G. Hanna et al. (Ed.), *Explanation and proof in mathematics: Philosophical and educational perspectives* (pp. 115–135). Springer Science + Business Media.

Blum, W., & Kirsch, A. (1991). Preformal proving: Examples and reflections. *Educational Studies in Mathematics, 22*(2), 183–203.

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

Dörfler, W. (2006). Diagramme und Mathematikunterricht [Diagrams and the mathematics classroom]. *Journal für Mathematikdidaktik, 27*(3/4), 200–219.

Dreyfus, T., Nardi, E., & Leikin, R. (2012). Forms of proof and proving in the classroom. In G. Hanna & M. de Villiers (Eds.), *Proof and proving in mathematics education. New ICMI study series* (Vol. 15). Springer Science+Business Media.

Gholamazad, S. (2007). Pre-service elementary school teachers' experiences with the process of creating proofs. In J. Woo, H. Lew, K. Park, & D. Seo (Eds.), *Proceedings of the 31th Conference of the International Group for the Psychology of Mathematics Education* (pp. 265–272). Seoul: PME.

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.

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education, 31*(4), 396–428.

Hemmi, K. (2008). Students' encounter with proof: The condition of transparency. *ZDM—The International Journal on Mathematics Education, 40,* 413–426.

Knuth, E., Choppin, J., & Bieda, K. (2009). Middle school students' production of mathematical justifications. In D. Stylianou, M. Blanton, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. 153–170). New York: Routledge.

Krummheuer, G. (1999). The narrative character of argumentative mathematics classroom interaction in primary education. In *Proceedings of the European Society for Research in Mathematics Education I* (pp. 331–341). Osnabrück: Forschungsinstitut für Didaktik der Mathematik.

Krummheuer, G. (2013). The relationship between diagrammatic argumentation and narrative argumentation in the context of the development of mathematical thinking in the early years. *Educational Studies in Mathematics, 84*(2), 249–265.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.

Lekaus, S., & Askevold, G.-A. (2015). Dialogues as in instrument in mathematical reasoning. In B. Di Paola & C. Sabena (Eds.), *Teaching and learning mathematics: Resources and obstacles, Proceedings of CIEAEM 67* (pp. 399–403). Aosta, Italy.

Mellin-Olsen, S. (1984). *Eleven, matematikken og samfunnet*. Oslo: NKI-forlag.

Mellin-Olsen, S. (1996). Oppgavediskursen i matematikk. *Tangenten, 7*(2), 9–15.

Ministry of Education and Research. (2013). *Curriculum for the common core subject of mathematics* (*MAT1-04*). Oslo: Norwegian Directorate for Education and Training. https://www.udir.no/kl06/MAT1-04?lplang=eng. Accessed: March 08, 2017.

Peirce, C. S. (1978). *Collected papers of Charles Sanders Peirce*. Cambridge, MA: Harvard University Press.

Skemp, R. R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching, 77,* 20–26.

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

Wenger, E. (1998). *Communities of practice*. Cambridge: Cambridge University Press.

Wille, A. (2011). Activation of inner mathematical discourses of students about fractions with the help of imaginary dialogues: A case study. In *Proceedings of the 35th Conference of the International Group for the Psychology of Mathematics Education* (pp. 337–344). Ankara, Turkey: PME.

Wille, A. (2013). Mathematik beim Schreiben denken—Auseinandersetzung mit Mathematik in Form von selbst erdachten Dialogen. In M. Rathgeb, M. Helmerich, R. Krömer, K. Lengnink, & G. Nickel (Eds.), *Mathematik im prozess* (pp. 239–254). Wiesbaden: Springer Fachmedien.

Wille, A., & Boquet, M. (2009). Imaginary dialogues written by low-achieving students about origami: A case study. In *Proceedings of the 33rd Conference of the International Group for the Psychology of Mathematical Education* (Vol. 1, pp. 337–344). Thessaloniki, Greece: PME.

Chapter 12
The Need for "Linearity" of Deductive Logic: An Examination of Expert and Novice Proving Processes

Shiv Smith Karunakaran

Abstract Mathematicians have long claimed that the proving process cannot be considered a "linear" process and that undergraduates may view the proving process to be necessarily "linear". However, there is little empirical research that supports this familiar claim. Using grounded theory methods, expert and novice provers of mathematics were examined in the process of proving novel mathematical statements. Expert provers of mathematics were willing to knowingly and temporarily interrupt the deductive logic of their proving process in order to make progress towards constructing an eventually complete deductive argument. On the other hand, novice provers seemed less inclined to behave in a similar manner. They seemed to rigidly require the deductive logic of the proving process to remain intact. In light of these findings, implications for mathematics curricula writers and mathematics instructors are discussed.

Keywords Proof · Proving · Mathematical argumentation · Expert–novice
Deductive thinking

Research has demonstrated that a gap may exist between a novice's and an expert's understanding of the proving process. Raman (2002) posited that a mere presentation of the statement of a theorem and subsequent presentation of the proof may not engender students' understanding of the proving process. This assertion is further supported by Chin and Tall's (2002) claim that mathematics textbooks present the process of a mathematical proof "as the development of a sequence of statements using only definitions and preceding results, such as deductions, axioms, or theorems" (p. 213). This systematic and step-by-step manner expressed in final written proofs does not reveal what led the mathematicians to produce the particular arguments. This may, in turn, promote the memorization of the proofs by the students without understanding the proving process. Students may look at proof as

S. S. Karunakaran (✉)
Michigan State University, East Lansing, USA
e-mail: karunak3@msu.edu

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_12

a finished product generated by someone else and not as a constructed argument that the students themselves could produce.

In fact, Romberg (1992) suggested that mathematics should no longer be thought of as a finished product, but as a "process of inquiry and coming to know, a continually expanding field of human creation and invention" (p. 751). This revised view of mathematics focuses on the perspective that to know mathematics is to do mathematics. This, in turn, calls for more focus on the mathematical work and activity of practicing mathematicians, which could subsequently serve as a model for students' mathematical activity. Lampert (1990) and Stylianou (2002) have called for one of the goals of mathematics instruction to better understand the practices of mathematicians. Furthermore, Weber and Mejia-Ramos (2011) have also listed as one of the goals of mathematics instruction to be for students to behave more like mathematicians in proof-related activities or tasks. In other words, it is important to better understand authentic mathematical practice (as engaged in by experts), and to use that understanding to realize what authentic mathematical experiences may be for students. To work towards these goals, researchers (Blum and Kirsch 1991; Weber 2001) have emphasized how teachers can help students better learn from and understand a mathematician's work with proof by making the act or process of proving clearer to the students. Also, Stylianides (2007) has emphasized how the notion of proof in all grades should both be true to the nature of mathematics as a discipline and honor the student as an authentic mathematical learner. To this end, the study from which this chapter is derived aimed to contribute new knowledge about the similarities and differences between the observed use of mathematical knowledge by expert provers and novice provers while proving a mathematical statement.

The Proving Process

Boero (1999) wrote about the general "phases" involved within the process of proving a conjecture or a mathematical statement. These phases can be described as follows:

1. Exploration of the content (and limits of validity) of the conjecture; heuristic, semantic (or even formal) elaborations about the links between hypotheses and thesis; identification of appropriate arguments for validation, related to the reference theory, and envisaging of possible links amongst them.
2. Selection and enchaining of coherent, theoretical arguments into a deductive chain, frequently under the guidance of analogy or in appropriate, specific cases, etc.
3. Organization of the enchained arguments into a proof that is acceptable according to current mathematical standards.
4. Approaching a formal proof (Boero 1999).

Boero also distinguished between mathematical proof as a product, and proving as the process by which a proof may be generated. He emphasized that these phases cannot be separated and that they are not used linearly in the process of proving by mathematicians. As described by Boero (1999), rather than only following a straight logical argument structure, mathematicians continuously move between exploratory, inductive, and deductive processes.

Carlson and Bloom (2005) described the more general process of problem solving in terms of a framework that has four phases. Based on interviews with 12 mathematicians, Carlson and Bloom (2005) developed a “problem-solving framework” that has four phases: orientation, planning, executing, and checking. Carlson and Bloom’s four phases seem loosely aligned to those of Boero. During the orienting phase, the problem-solver is involved in the initial engagement with the problem statement or task. The next phase of planning involves the solver developing initial conjectures of how to arrive at a viable solution. The phase of executing involves carrying out the strategies developed in the planning phase, while the checking phase involves the verification of the validity of the solution generated.

In both of these frameworks, of Boero (1999) and of Carlson and Bloom (2005), as the provers/problem-solvers decide on the viability of their strategies, the *conjecture–imagine–evaluate* subcycle was repeated until a viable solution path was identified. Unfortunately not much is known about how provers of varying levels of mathematical expertise attempt to construct a proof. As a step towards addressing this need for research, this study sought to compare aspects of how graduate students of mathematics (considered to be expert provers) and undergraduate students of mathematics (considered to be novice provers) organize different phases in the service of proving mathematical statements. Although this study does not explicitly use the phases of proving or problem solving, as described by Boero (1999) and Carlson and Bloom (2005) respectively, these phases describe how the processes of proving and problem solving involve similar themes that highlight how theses phases can be nonlinear and iterative in nature.

Use of Mathematical Knowledge in Proving

Gaining expertise in the act of doing mathematics, and thus in the process of proving mathematical statements (which is a subset of doing mathematics), involves the use of the existing knowledge that the individual has accumulated. It is not merely the fact that individuals have assimilated this wealth of knowledge, but also how they call upon the various facets and parts of this knowledge that allows them to demonstrate expertise in the act of proving mathematical statements. To this end, it seems useful to examine portions of the mathematical knowledge that can be inferred as an individual proves a mathematical statement. These inferences about mathematical knowledge can be drawn from observations of the properties, objects, procedures, definitions, theorems, and so on, that an individual brings to bear in the service of proving a statement. The parts of mathematical knowledge that are called on during the process of proving are being referred to in this chapter as *resources*. It is further posited that an individual's mathematical knowledge is comprised of a connected network of such resources. This view of mathematical knowledge and of doing mathematics as using a network of relations is not novel within mathematics education research (Hiebert and Lefevre 1986).

Once the individual is observed calling on one or more resources, then he or she can be reasonably observed acting on these resources. That is, he or she may use the resources to perform certain *actions* such as asking a question based on the resource (s), constructing an example, searching for a counterexample, and using a form of reasoning. The terms "actions" and "resources" are adapted and expanded from the work of Wilkerson-Jerde and Wilensky (2011). Those researchers described how mathematicians use different *resources of mathematical understanding* and *acts of mathematical understanding* in order to read and understand a published, but unfamiliar mathematical paper about knot theory. Wilkerson–Jerde and Wilensky do not offer any definition of resources of mathematical understanding other than to equate them to "specific knowledge" (2011, p. 22) that the mathematicians used in order to understand the unfamiliar mathematics present in the paper. The concept of acts of mathematical understanding has been expanded for the purposes of the present study to include a wider range of actions that the prover may utilize in the process of proving.

However, simply identifying the resources and the actions an individual uses in the process of proving does not give sufficient insight into the individual's rationale underlying his or her use of the resources and actions. Skemp's (1976) construct of relational understanding of mathematics involves knowing not just what to do when doing mathematics, but also why to do it. The assumption both Skemp and I make is that individuals make intentional decisions to use what they know and how to use what they know. Skovsmose (2005) differentiates the notions of *action* and *blind activity* using this same assumption. Blind activity is characterized by automatic behavior and it assumes that there is no true rationale behind what an individual is doing. In contrast, an action presupposes some degree of choice and as such assumes that the individual has a purposeful intention behind performing the said action. For the present study, I adopt Skovsmose's notion that any action or set of actions identified during the process of proving a statement is associated with intention(s). That is, one cannot truly describe the actions of an individual without considering the intention behind the actions.

Skovsmose (2005) hints towards going beyond merely identifying singular actions and the attached intentions. He seems to suggest that it is also important to identify *activities* (or actions grouped together, as defined by Skovsmose) and the intentions behind such activities. Thus, when analyzing an individual's process of proving, I also went beyond identifying individual actions (and the resources involved) to identify groups of actions and resources that seemed to be tied together with a common intention or intentions. Analogously to Skovsmose's notion of activities, I defined these groups of actions and resources as *bundles*. More specifically, bundles are defined as subsections of the proving process that consist of groups of actions and resources that are clustered together by identifiable intentions. These identifiable intentions are nested within the assumed larger goal of proving the statement in question.

The theoretical constructs of bundles, the associated intentions, and the constituent actions and resources were used to describe an individual’s (either an expert or a novice) use of mathematical knowledge in their dynamic proving process. The research questions that guided the larger study (Karunakaran 2014) dealt with the examination of the ways in which expert and novice provers’ use of bundles and their associated intentions were similar and/or different. The specific research question reported on here is: *How do expert and novice provers of mathematics sequence their logic in the service of proving novel mathematical statements?*

Methods

The research question described previously does not fall under the category of validating an existing theory of how individuals prove. Instead the research was about investigating the bundles (and the component resources and actions) and their associated intentions of different groups of individuals (expert and novice provers of mathematical statements) in the process of proving a mathematical statement. The focus of the research lent itself to the adoption of certain *grounded theory* methods to guide data collection and data analysis, specifically the data analysis strategies of *open coding, axial coding, selective coding,* and *constant comparative analysis* (Charmaz 2006; Strauss and Corbin 2008).

A group of novice provers of mathematics and a relatively more expert group were recruited. The novice group included five undergraduate students who had all successfully completed at least one proof-based course in real analysis. Placing such a requirement allowed the researchers to be confident that the members of the novice group (hereby referred to as Novice Provers or NPs) had been exposed to an introduction to various mathematical proof strategies, such as proof by mathematical induction, proof by contradiction, and proof by first principles. The expert group included five doctoral students in mathematics, all of whom had successfully passed their department’s doctoral qualifying examinations. By requiring that the members of the expert group (hereby referred to as Expert Provers or EPs) to have passed doctoral qualifying examinations ensured their experience with doing mathematics, and more specifically in proving and in the generation of proofs.

All ten participants (NPs and EPs together) were each presented with five real analysis statements (see Fig. 12.1) in an interview setting using a think-aloud protocol. These interviews were video-recorded, and then were fully transcribed. As is shown in the figure, the directions for all five tasks were to “Validate or refute the following statement.” Presenting this as the direction, and not presenting the more traditional direction of “Prove that …”, ensured that the initial opinion of the participants about the truth of the mathematical statement was not merely due to the format of the directions of the task statement. Also, the statements of Tasks 1 and 4

1. Validate or refute the following statement:
   Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of real numbers such that $0 < a_n \leq a_{2n} + a_{2n+1}$, $\forall\, n \in \mathbb{Z}\ \&\ n \geq 1$. Then the series $\sum_{n=1}^{\infty} a_n$ converges.

2. Validate or refute the following statement:
   There exists a function g that satisfies all of the following properties:

   (a) $g(x) = x \cdot g\left(\frac{1}{x}\right)$, for all real numbers $x \neq 0$, and

   (b) $g(x) + g(y) = 1 + g(x + y)$, for all real numbers x and y.

3. Validate or refute the following statement:
   Let $f(x) = \cos(\sqrt{1}x) + \cos(\sqrt{2}x) + \cos(\sqrt{3}x) + \cdots + \cos(\sqrt{n}x)$, where $n \in \mathbb{Z}^{+}$. Then $f(x)$ is a periodic function for only $n = 1$.

4. Validate or refute the following statement:
   Suppose $p(x)$ is a polynomial with integer coefficients and such that $p(x)=3$ has three distinct integer solutions. Then the equation $p(x)=4$ has at least one integer solution.

5. Validate or refute the following statement:
   Suppose $p(x)$ is a polynomial with integer coefficients such that the sum of all of its coefficients is an odd integer, and its constant coefficient is also odd. Then there is no such polynomial with at least one integer root.

Fig. 12.1 The five tasks presented to both the expert and novice provers

are false as stated in Fig. 12.1. If the provers came up with a valid proof of why the statements were false, then the statements were amended by the researcher to make them true, and the provers were again asked to validate or refute the amended statement.

Results

Only one of the claims generated by the findings of the larger study is reported here. This claim involves how the expert and novice provers seemingly differ in their organization and utilization of deductive logic within the process of proving. The assumed larger goal for each of the participants was to construct a deductive argument that validates or refutes the mathematical statement, and as such, it was expected that both the expert and novice provers would seek to find a sequence of bundles that when chained together would produce a complete deductive argument that validates or refutes the mathematical statement. Such a complete deductive argument should contain no interruptions in the deductive logic from one bundle to the next. In other words, a deductive rationale could be identified for every step of the final deductive argument generated. And as was expected, both the expert and novice provers sought to generate such a complete and uninterrupted deductive sequence of bundles, especially when they were observed to easily move through the proving process without getting “stuck” anywhere.

An instance of a prover making efficient progress or progress in which the prover was not observed getting “stuck” occurred with Kevin (NP) as he made progress through Task 2 (see Fig. 12.1 for the statement of Task 2). In his work on Task 2, Kevin was initially observed working with a bundle in which he looked for a contradiction between the two given properties of the function $g(x)$. He further validated this by stating his intention, “I was kind of hoping to get to a nice contradiction to show that there didn’t exist a function that could satisfy all these statements.” He did this by trying to symbolically manipulate the two equations given in the properties. However, he did not immediately arrive at an obvious contradiction. At this point, his intention seemed to change because he noticed a pattern in the three values he generated for the function (i.e., $g(0) = 1$, $g(\sqrt{2}) = \frac{1}{1-\sqrt{2}}$, and $g(-\sqrt{2}) = \frac{1}{1+\sqrt{2}}$). This indicated a change in bundles and the new bundle seemed to have the new intention of generating a function that is consistent with the three values generated for the function. Kevin initially made a prediction that the function $g(x) = \frac{1}{x-1}$ might satisfy the given properties in Task 2. However, he quickly refuted this claim. His second prediction, $g(x) = x + 1$, was consistent with all three values that he generated. Once again, his intention seemed to change to now wanting to verify the validity of the function $g(x) = x + 1$ with the given properties, and this corresponded to a change in his work to a bundle in which he verified that his predicted function was consistent with the two properties. This sequence of three bundles (finding a contradiction → predicting a function consistent with generated values → verifying the validity of $g(x) = x + 1$) has no interruptions in the deductive logic. That is, there were no instances in this sequence of bundles in which he needed to assume the deductive validity of any claim.

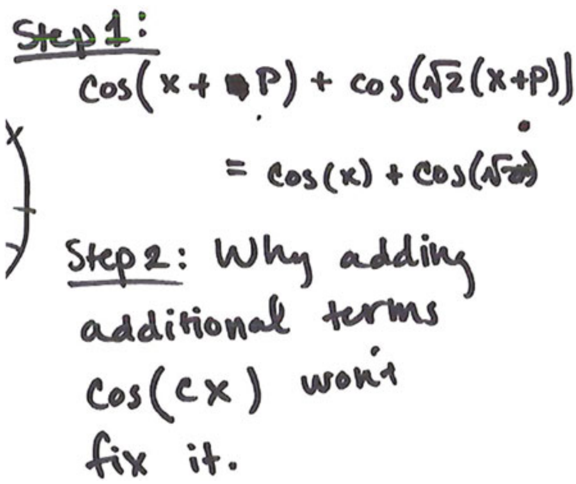
Kevin’s (NP) behavior of seeking a sequence with no interruptions in the deductive logic was not unique to him. The other novice provers and the expert provers also were observed with the same behavior, as long as the prover seemed to make quick progress with the proving of the statement. However, when the provers seemed to struggle more with finding such a sequence of bundles, the expert provers demonstrated that they were willing to include interruptions in the deductive logic as they sequenced bundles, and as such, were more flexible in their sequencing of bundles. Moreover, the expert provers seemed willing to consider sequencing bundles together even when they had only an intuitive rationale for doing so. For instance, when working on Task 3 (see Fig. 12.1 for the statement of Task 3), Julie (EP) reasoned validly that the function $f(x) = \cos(\sqrt{1}x) + \cos(\sqrt{2}x) + \cdots + \cos(\sqrt{n}x)$ is periodic for $n = 1$. She then examined the function for $n = 2$ (i.e., $f(x) = \cos(\sqrt{1}x) + \cos(\sqrt{2}x)$) and stated that the summands $\cos(\sqrt{1}x)$ and

$\cos(\sqrt{2}x)$ have periods 2π and $\sqrt{2}\pi$ respectively. She then made a claim that the function $f(x) = \cos(\sqrt{1}x) + \cos(\sqrt{2}x)$ cannot be periodic because the periods of $\cos(\sqrt{1}x)$ and $\cos(\sqrt{2}x)$ are not rational multiples of each other. She explained that this latest claim was unsubstantiated, but she was going to assume that it was valid in order to make progress with the proving process. In fact, she broke the proof of the statement in Task 3 as being comprised of two steps (see Fig. 12.2). She then stated, "I don't know how to prove this first step." She went on to prove the second step, and arrived at a finished, albeit not fully justified argument.

Bartok (EP) also illustrated flexibility in the sequencing of bundles in his work with Task 1 (see Fig. 12.1 for the statement of Task 1). After Bartok correctly generated a counterexample for the original statement for Task 1, he was presented with the modified statement that asked to validate or refute that the series $\sum_{n=1}^{\infty} a_n$ always diverges. When considering how to approach the modified statement, Bartok brought up the harmonic series. He quickly concluded that the harmonic series is not an example of the type of series being considered in the modified Task 1, since the harmonic series does not satisfy the inequality condition present in the modified statement. He also made an additional observation that the "harmonic [series] doesn't quite get up to [the inequality] condition." When asked to further explain his observation, he proceeded to explain that for the harmonic series, $\frac{1}{2n} + \frac{1}{2n+1} < \frac{1}{n}$, but "only barely". For example, for $n = 1$, the left side of the inequality is $\frac{1}{2} + \frac{1}{3} = \frac{5}{6}$, which is just "barely" less than 1. This led Bartok to make the following statement:

> Um, which means that… if [the terms of the harmonic series] were a bit bigger it would satisfy [the inequality in the statement of Task 1] which means that… this would say that any thing that satisfies this would be slightly bigger than the harmonic series.

Fig. 12.2 Julie's written out steps outlining her argument validating the statement in Task 3



Bartok (EP) was asked to explain why he thought that any series that would satisfy the inequality condition would be "slightly bigger" than the harmonic series. He made it clear that he did not have any "mathematical" or deductive rationale for the claim, but that he intuitively believed this to be true. At this point, he stated that the modified statement was true. He elaborated on his conclusion by stating that since the terms of any series that satisfied the inequality condition would be bigger than the terms of the harmonic series, and since the harmonic series is known to be a diverging series, that would mean that the series $\sum_{n=1}^{\infty} a_n$ would also diverge (using the Direct Comparison Test for Divergence). However, he was clear that he would need to "formally prove" his intuitive claim. The rest of Bartok's work for Task 1 involved trying to prove that for any sequence $\{a_n\}$ that satisfies the inequality condition, $a_n \geq \frac{1}{n}, \forall n$. In this instance, Bartok was willing to interrupt the deductive logic in his sequencing of bundles by allowing himself to intuitively accept the truth of the claim that the terms of any series satisfying the inequality condition in Task 1 needs to correspondingly be bigger than the terms of the harmonic series.

When expert provers are faced with a question regarding the validity of a claim (e.g., Bartok's intuitive claim or Julie's unsubstantiated claim) that needs to be resolved for the proving process to move forward without any interruptions in the deductive sequencing of bundles, it seems to be acceptable to the expert provers to assume the validity of the claim in order to make further progress. They seem not to have the requirement that every step of the proving process has to be mathematically resolved in a linear fashion. They seem capable of setting aside a question temporarily, even though they clearly recognize that the question or claim needs to eventually be resolved completely. In other words, the expert provers seem to allow for interruptions (namely, setting aside a deductively unresolved question temporarily) within a deductive sequence of bundles (see Fig. 12.3).

This is in contrast to the work of most of the novice provers who seem to require all the questions to be fully answered (have no interruptions in the deductive logic in the sequencing of bundles) in order to make progress with the proving process (see Fig. 12.4).

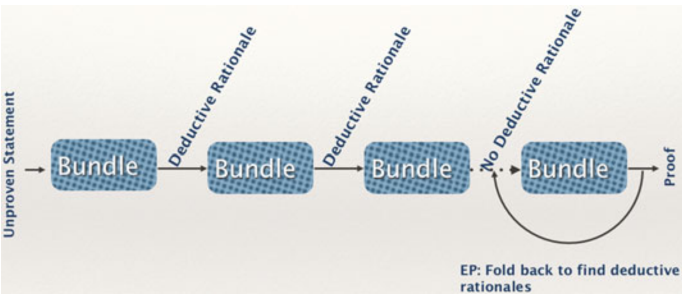


Fig. 12.3 Expert provers' sequencing of bundles in their proving process

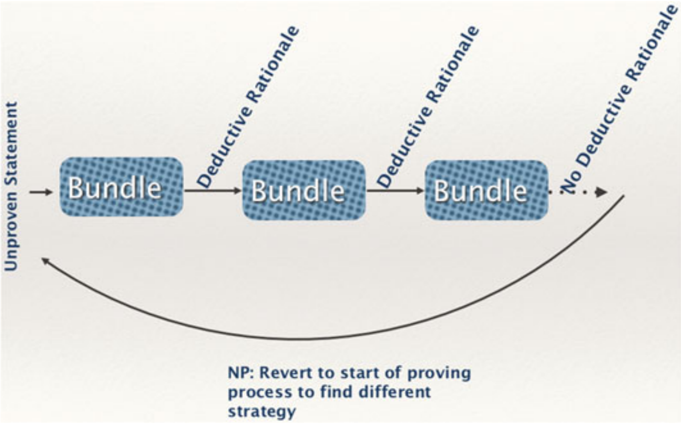


Fig. 12.4 Novice provers' sequencing of bundles in their proving process

Discussion

The results described here focus on how expert provers may use intuition in order to assume the answers to questions that are yet unresolved, in order to move on with their proving process. The expert provers also are aware that the interruptions in the deductive logic of the sequencing of bundles (perhaps by using intuition to validate claims) need to eventually be deductively resolved for them to successfully produce a proof. So, if the sequence of bundles generated by allowing for interruptions leads the expert prover to an argument that validates or refutes the mathematical statements, he or she knows that he or she will have to return to each one of those interruptions to try to produce a deductive argument to resolve them. Thus, expert provers seem to be aware that they may need to move back and forth between allowing interruptions and then having to return later to deductively resolve those interruptions. In contrast, the novice provers in this study seem to require sequences of bundles to have uninterrupted deductive logic to be able to move on with the proving process. The novice provers' proving process does seem to be consistent with what Boero (1999) called the linear model of proof and with what Reiss and Renkl (2002) described to be the false conception held by students about the linearity of proof. Boero (1999) wrote that a linear model of proof involves just a straight logical argument structure and does not involve transitioning back and forth between exploratory, inductive, and deductive processes. Reiss and Renkl (2002) described students having a false conception that a "proof is a straight-forward, deductively deduced, systematic and logical sequence of steps" (p. 32). Novice provers in this study acted in a manner that suggests that they could have possessed this false conception of proof. As such, this study provides empirical evidence to support the claim that expert and novice provers approach the sequencing of bundles in different ways.

At the outset of attempting to prove the mathematical statements, expert provers in this study were observed being comfortable using intuitive reasoning to resolve questions that may arise while moving from one bundle to the next. This behavior is not completely explained by other research in expert proof production and proving (Hanna and Barbeau 2008; Rav 1999; Weber 2001, 2004; Weber and Alcock 2004, 2009). Weber and Alcock (2004) distinguished between syntactic and semantic proof productions. Syntactic proof productions are characterized solely by the use of correctly stated definitions and other facts present in the mathematical statement in order to produce a proof. Semantic proof productions are characterized by the use of examples, diagrams, and other instantiations to suggest more formal methods to produce the proofs. These two forms of proof productions also seem consistent with the two forms of proof described by Rav (1999). Rav described two forms of proof: derivations or formal proofs, and conceptual or informal proofs. Derivations or formal proofs are analogous to Weber and Alcock's syntactic proof productions and are described as "syntactic objects of some formal system" (Rav 1999, p. 11). These proofs are strictly bound by the rules of logical inference. On the other hand, conceptual proofs or informal proofs are akin to Weber and Alcock's (2004) semantic proof productions and are rigorous arguments that are deductively sound, but may not contain precise mathematical definitions. Rav (1999) posits that mathematicians readily accept conceptual proofs and this notion is consistent with Weber and Alcock (2004). Rav describes conceptual proofs to be more conducive to understanding the structure of why mathematical statements are true. The behavior of the expert provers described here does not seem consistent with either semantic proof productions or conceptual proofs. These analogous forms of proofs both require a deductive argument structure. The proving process that was followed by expert provers in this study was characterized by allowing for intuitively resolved interruptions in the deductive logic in order to move from one bundle to the next.

The results reported here has implications for the developers of mathematics curricula. The display and treatment of proofs in many current undergraduate mathematics textbooks have tended towards presenting both the proving process and proofs in a strictly sequential or linear style (Alibert and Thomas 1991). The current study offers evidence of how expert provers do not engage in the proving process in a strictly linear or sequential manner. As an implication of these results, undergraduate curriculum developers are presented with an argument for the presentation of proofs in textbooks to reflect the nonlinear manner in which the proving process can occur and convey to students the idea that it is acceptable to assume the truth of a statement during the proving process as long as one eventually returns to address that assumption. Moreover, this argument extends to the curricula of lower grades. In fact, Wong and Sutherland (this volume) demonstrate that a common curriculum used in grades 10–12 in Hong Kong could benefit from a more in-depth presentation of proof-related examples.

In a similar manner, this study has implications for undergraduate mathematics instructors who teach proof and proving. Just like the presentation in the textbooks, undergraduate mathematics instructors may present proving as a sequentially derived series of steps. This presentation may lead to students having very little understanding of how the proof was developed (Raman 2002). More recent research echoes the need for more coherent presentation of proofs during undergraduate instruction. Gabel and Dreyfus (2017) discuss the use of an intervention designed to help an undergraduate mathematics instructor to present a proof as more of a coherent story, and to more completely link the different parts of the proof. Also, Tsujiyama and Yui (this volume) discuss how merely presenting accurate, complete, and successful arguments to students will not necessarily deepen their understanding of the argument. In fact, the presentation and examination of unsuccessful arguments allowed students to be able to compare and contrast the successful and unsuccessful arguments (Tsujiyama and Yui, this volume). This in turn offered the students a context in which they could explore why the successful argument worked, and why the unsuccessful argument failed (Tsujiyama and Yui, this volume).

References

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

Blum, W., & Kirsch, A. (1991). Preformal proving: Examples and reflections. *Educational Studies in Mathematics, 22,* 183–203.

Boero, P. (1999). Argumentation and mathematical proof: A complex, productive, unavoidable relationship in mathematics and mathematics education. *International Newsletter on the Teaching and Learning of Mathematical Proof*. Retrieved from http://www.lettredelapreuve.it/OldPreuve/Newsletter/990708Theme/990708ThemeUK.html.

Carlson, M., & Bloom, I. (2005). The cyclic nature of problem solving: An emergent problem-solving framework. *Educational Studies in Mathematics, 58,* 45–75.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative research*. Thousand Oaks, CA: SAGE Publications.

Chin, E. -T., & Tall, D. (2002). Proof as a formal procept in advanced mathematical thinking. In F. -L. Lin (Ed.), *Proceedings of the International Conference on Mathematics: Understanding Proving and Proving to Understand* (pp. 212–221). Taipei, Taiwan: National Taiwan Normal University.

Gabel, M., & Dreyfus, T. (2017). Affecting the flow of a proof by creating presence—A case study in Number Theory. *Educational studies in mathematics*. https://doi.org/10.1007/s10649-016-9746-z.

Hanna, G., & Barbeau, E. (2008). Proof as bearers of mathematical knowledge. *ZDM Mathematics Education, 40,* 345–353.

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1–28). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Karunakaran, S. (2014). *Comparing bundles and associated intentions of expert and novice provers during the process of proving*. Unpublished doctoral dissertation. The Pennsylvania State University, University Park, PA.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal, 27,* 29–63. https://doi.org/10.2307/1163068.

Raman, M. J. (2002). *Proof and justification in collegiate calculus*. Unpublished doctoral dissertation. University of California, Berkeley, CA.

Rav, Y. (1999). Why do we prove theorems? *Philosophia Mathematica, 7*(1), 5–41.

Reiss, K., & Renkl, A. (2002). Learning to prove: The idea of heuristic examples. *Zentralblatt für Didaktik der Mathematik (ZDM), 34*(1), 29–35. https://doi.org/10.1007/BF02655690Raman.

Romberg, T. A. (1992). Problematic features of the school mathematics curriculum. In P. W. Jackson (Ed.), *Handbook of research on curriculum: A project of the American Educational Research Association*. New York, NY: Macmillan.

Skemp, R. R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching, 77,* 20–26.

Skovsmose, O. (2005). Meaning in mathematics education. In J. Kilpatrick, C. Hoyles, O. Skovsmose, & P. Valero (Eds.), *Meaning in mathematics education* (pp. 83–100). New York, NY: Springer.

Strauss, J., & Corbin, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.) [Kindle edition]. Los Angeles, CA: SAGE Publications.

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

Stylianou, D. A. (2002). On the interaction of visualization and analysis: The negotiation of a visual representation in expert problem solving. *Journal of Mathematical Behavior, 21,* 303–317.

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

Weber, K. (2004). Traditional instruction in advanced mathematics courses: A case study of one professor's lectures and proofs in an introductory real analysis course. *Journal of Mathematical Behavior, 23,* 115–133.

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

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

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

Wilkerson-Jerde, M. H., & Wilensky, U. J. (2011). How do mathematicians learn math? Resources and acts for constructing and understanding mathematics. *Educational Studies in Mathematics, 78,* 21–43.

Chapter 13
Reasoning-and-Proving in Algebra in School Mathematics Textbooks in Hong Kong

Kwong-Cheong Wong and Rosamund Sutherland

Abstract To promote learning mathematics with understanding, mathematics educators in many countries recommend that proof (and proof-related reasoning) should play a central role in school mathematics. In response to this recommendation, this study examines the opportunities for students to learn reasoning-and-proving from solving algebra problems in a popular school mathematics textbook from Hong Kong. The study adopts the methodology of Stylianides (2009). Results show that such opportunities are relatively limited. Furthermore, the overwhelming majority of the demonstration proofs require little reasoning. There are almost no opportunities for conjecturing, but many instances of empirical non-proof arguments. Overall, the results suggest that proof plays a marginal role in school mathematics in Hong Kong.

Keywords Reasoning-and-proving · School mathematics textbooks
Algebra · Hong Kong

Introduction

Besides verifying the truth of a mathematical statement, proof can have many other functions in mathematics, including explanation, which can promote sense making and understanding in mathematics (de Villiers 1990). As a consequence, many mathematics educators around the world, especially those in the U.S., recommend that proof and proof-related reasoning permeate school mathematics across all grade levels and content areas (e.g., Ball et al. 2002; NCTM 2000; Stylianou et al. 2010; Hanna and de Villiers 2012). Furthermore, since textbooks can have an

K.-C. Wong (✉)
The Hong Kong Polytechnic University, Hung Hom, Hong Kong
e-mail: wongkwongcheong@gmail.com

R. Sutherland
University of Bristol, Bristol, UK

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_13

influence on what the student learns, many studies have been conducted in different countries to examine the opportunities for students to learn reasoning and proof from school mathematics textbooks (e.g., Hanna and de Bruyn 1999; Nordström and Löfwall 2005; Stacey and Vincent 2009; Stylianides 2009; Thompson et al. 2012; Davis 2012; Bieda et al. 2014; Davis et al. 2014; Fujita and Jones 2014; Otten et al. 2014a, b; Hunte 2016). However, almost all of these studies were conducted in Western countries and few have been conducted in East Asian countries (e.g., Singapore, South Korea, Hong Kong) where students have consistently performed very well in international studies of mathematics achievement such as TIMSS (e.g., Mullis et al. 2012). The present study aims to complement the research knowledge of the field by examining the opportunities for students to learn reasoning and proof when they are solving algebra problems in a popular secondary school mathematics textbook from Hong Kong. It is expected that the results obtained will shed light on how reasoning and proof is being treated in school mathematics in one of those high-achieving countries (or regions) and provide insights into the influences that Chinese culture may have on issues concerning understanding in school mathematics.

Literature Review

As mentioned above, in recent years many studies have been conducted in different countries to examine the opportunities for students to learn reasoning and proof (RP) from school mathematics textbooks. In this section we briefly review this body of literature, which can roughly be classified into three phases according to the sophistication of the analytic frameworks used. In the first phase, the analytic frameworks used were rather primitive. This phase started with Hanna and de Bruyn (1999), who examined two popular Ontarian grade-12 advanced-level mathematics textbooks, in the topics of algebra, geometry, functions and relations, exponents and logarithms, and trigonometry, using a framework consisting of three categories: proof, discussion of proof, and non-proof. They found that one textbook had 21% of its expository items and 21% of its exercises involving proof or discussion of proof compared to 17% and 16%, respectively, in the other. In Sweden, Nordström and Löfwall (2005) investigated proof in two popular Swedish upper secondary mathematics textbooks, in the topics of algebra, geometry, statistics and probability, functions and calculus, exponents and logarithms, trigonometry and complex numbers. They found that even in geometry "the occurrence of proof or discussion of proof was very low compared to the Ontarian textbooks" (p. 451). In Australia, Stacey and Vincent (2009) examined the modes of reasoning used in seven topics in the exposition sections of nine Australian grade-8 mathematics textbooks. They found a total of seven modes of reasoning: deduction using a general case, deduction using a specific case, deduction using a model, concordance of a rule with a model, experimental demonstration, appeal to authority, and qualitative analogy. They also found that "deductive explanations were common for some topics (especially multiplication of negative integers and area of a trapezium [both 100%]) but infrequent for others (especially division of fractions [only 17%])" (p. 285).

In the second phase, two more sophisticated frameworks were proposed based on their respective conceptualizations of proof-related constructs. This phase started with Stylianides (2009), who gave a conceptualization of *reasoning-and-proving*- (RP) to encompass four major proof-related activities: identifying patterns (definite and plausible), making conjectures, providing proofs, and providing non-proof arguments. Based on this conceptualization, he proposed an analytical framework to examine the RP opportunities in a reform-oriented grades 6–8 textbook series in the topics of algebra, number theory and geometry. He found that about 40% of the examined tasks were RP tasks and about 5% were proof-providing tasks. He found also that 97% of the plausible patterns and 88% of the definite patterns did not lead to conjectures, and 70% of the conjectures did not lead to proofs. Thompson et al. (2012) proposed a conceptualization of *proof-related reasoning* (RP) and developed an analytic framework for textbook analysis, which was used to examine the RP opportunities in twenty U.S. contemporary high school mathematics textbooks in the topics of exponents, logarithms and polynomials. They found that 5.4% of the exercises contained RP.

In the third phase, the two influential analytical frameworks by Stylianides (2009) and Thompson et al. (2012) were adapted to conduct textbook analysis of RP opportunities. Using a framework adapted from Stylianides (2009), Davis (2012) examined the dispersion of RP involving polynomial functions set within three differently organized (conventional, hybrid and reform-oriented) secondary mathematics textbooks in the U.S., and found that they contained, respectively, 4%, 9%, and 22% RP instances within the exercise sections. In another study, Davis et al. (2014) examined RP opportunities in two U.S. reform-oriented secondary advanced algebra textbooks. They found that 10.80% were RP tasks within exercise sections and 24.75% were RP sentences within exposition sections in one textbook compared to 15.76% and 30.23%, respectively, in the other. They found also that patterns in both textbooks were rarely used to develop conjectures or valid arguments. Using a framework adapted from Thompson et al. (2012), Otten et al. (2014a, b) examined RP opportunities for students (ages 13–16) in six U.S. secondary geometry textbooks. They found that about 25% of the exercises were RP tasks and less than 5% were proof-constructing tasks. They found also that it was rare for the RP process itself to be an explicit object of reflection. Also in the U.S., Bieda et al. (2014), in order to complement the research on RP in secondary mathematics, analyzed seven grade-5 elementary mathematics textbooks for RP opportunities. They adapted their framework from Stylianides (2009) and Thompson et al. (2012), and found 3.7% of the examined tasks involving RP. In Japan, Fujita and Jones (2014) analyzed the geometry component of a popular grade-8 Japanese textbook, using a framework based mainly on the work associated with the Third International Mathematics and Science Study (TIMSS) and informed by the work of Stylianides (2009) and of Thompson et al. (2012). They found that 70% of the examined exercises involved direct proofs and 24% involved making conjectures before proving. More recently, Hunte (2016) employed the framework of Otten et al. (2014a) to examine the RP opportunities in six geometry topics (triangles, congruent triangles, similar triangles, Pythagoras' theorem, quadrilaterals and circles) in three secondary textbooks. He found that the three textbooks had, respectively, 32%, 35% and 77% RP tasks within their exercise sections.

Several observations can be made about the studies reviewed above. First, RP opportunities in different textbooks varied greatly—from 3.7% (Bieda et al. 2014) to 77% (Hunte 2016). Second, higher grade levels tended to have more RP opportunities. Third, geometry tended to have more RP opportunities than algebra. Finally, reform-oriented curricula (which are proof-favouring) tended to have more RP opportunities than conventional curricula. But as noted by Cai and Cirillo (2014) one should be cautious about making comparisons between the findings across these studies, since different studies used different frameworks and methods (e.g., sampling and unit of analysis).

Against this background, our study set out to examine the opportunities for students to learn RP from solving algebra problems in a popular school mathematics textbook from Hong Kong.

The Context: Hong Kong SAR

Hong Kong, being a Special Administrative Region (SAR) of China, has its own school curriculum, which is mandated by its Education Bureau and is different from that in China. In 2009, the Hong Kong SAR Government (HKSARG) launched its New Senior Secondary (NSS) Mathematics Curriculum for students of Years 10–12 (ages 15–17). This new curriculum consists of two parts: the Compulsory Part and the Extended Part, the latter being more advanced and optional (see Education Bureau HKSARG 2007). In the present study, we focused on the Compulsory Part of the curriculum, since this part is the common core that all students have to study. Comprising three strands ("Algebra and Number", "Measures, Shape and Space", and "Data Handling"), this Compulsory Part is aimed at developing in students: "(a) the ability to think critically and creatively, to conceptualize, inquire and reason mathematically, and to use mathematics to formulate and solve problems in daily life, as well as in mathematical contexts and other disciplines; (b) the ability to communicate with others and express their views clearly and logically in mathematical language; …" (ibid., p. 2). These aims are not in any way unique—similar aims can be found, for example, in the Common Core Standards for School Mathematics in the United States (CCSSI 2010). Among these aims, one can find some indications of reasoning and proof (e.g., "to reason mathematically" and "to communicate with others and express their views clearly and logically in mathematical language"). Schools in Hong Kong are free to choose textbooks from a wide range produced by commercial publishers, but usually choose textbooks from the recommended textbook list provided by the Education Bureau, because these textbooks are guaranteed to be fully aligned with the new curriculum. The textbook series chosen for this study is *New Century Mathematics* (*NCM*) (Leung et al. 2014–2016), which is on the recommended textbook list and is coauthored by a prominent mathematics educator. Based on anecdotal evidence (to the best of our knowledge, there is no publication on the market statistics), *NCM* is one of the most, if not the most, popular textbooks currently in use in Hong Kong. Our rationale for choosing to conduct a textbook analysis instead of other means to investigate students' opportunities to learn reasoning and proof is that teachers in Hong Kong use textbooks frequently in their everyday planning and instruction. This has been pointed out by, for example, Tam et al. (2014):

> To many teachers [in Hong Kong], textbooks, rather than the official curriculum and assessment guides, are the 'curriculum documents,' as they rely heavily on textbooks in their teaching. (p. 101)

About this point, Leung and Park (2002) also noted:

> It seems that the [system] in Hong Kong … [is] relying on a kind of division of labour where the role of curriculum developers and textbook writers is to make sure that good pedagogy is embedded in the curricula and the textbooks. The teacher's role then is to follow the curriculum faithfully, and make sure that students follow the suggested procedures. This will free up teachers' time for preparation and allow them to have a heavier teaching load. (p. 128)

Therefore, it is likely that the treatment of reasoning and proof in textbooks influences teacher's instructional decisions and thus their students' opportunities to learn reasoning and proof.

Analytic Framework and Method

There are different notions of proof in mathematics education research (see, e.g., Stylianides 2007; Reid and Knipping 2010). In this study, we followed the influential conceptualization of the notion of proof in the context of school mathematics by Stylianides (2007), which emphasizes the socio-cultural aspects of students' proving practice:

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

This conceptualization of proof (and proving) is compatible with the analytic framework by Stylianides (2009) that we adapted in this study. Our reason for adapting Stylianides' framework is that our research project, of which this study was part, has similar aims to his, namely, we both aimed to investigate the reasoning and proof opportunities designed for students in one strategically selected textbook series across different levels and different content areas. Stylianides' (2009) framework is based on his conceptualization of *reasoning-and-proving* (RP), a term he proposed to describe the overarching activity that encompasses all of the four major proof-related mathematical activities: (a) identifying patterns, (b) making conjectures, (c) providing proofs, and (d) providing non-proof arguments. As shown in Table 13.1, the first two activities are grouped into the category of making mathematical generalizations and the latter two into the category of providing support to mathematical claims. The idea behind this conceptualization is that making mathematical generalizations (*identifying a pattern* and *conjecturing*) and providing support to mathematical claims (*proving*) are two fundamental and interrelated aspects of doing mathematics (Boero et al. 2007). Furthermore, there are two kinds of pattern: plausible and definite; two kinds of proof: generic example and demonstration; and two kinds of non-proof argument: empirical argument and rationale. An important difference between Stylianides' original framework and our adapted version of his framework is that we use a broader category of Demonstration than his; specifically, we include in the category of Demonstration also "Proof by Definition" and "Proof by Calculation" (see Table 13.3) which are frequently used in the exercises of the textbook series we examined.

In this study, following Hanna and de Bruyn (1999), we focused on the Compulsory Part of the curriculum and examined all the algebraic chapters (i.e., chapters in the "Algebra and Number" strand) in Year 10 of the chosen textbook series. (For RP opportunities in the other two strands, see Wong 2017a, b.) This involved all of the eight chapters in Books 4A and 4B of *New Century Mathematics*: Chapter 1 Number Systems, Chapter 3 Quadratic Equations in One Unknown, Chapter 4 Basic Knowledge of Functions, Chapter 5 Quadratic Functions, Chapter 6 More about Polynomials, Chapter 7 Exponential Functions, Chapter 8 Logarithmic Functions, and Chapter 9 Rational Functions. Following Stylianides (2009), we focused on the exercise sections in these chapters. In each of these chapters, exercises were categorized under various headings: Q&A, Review Exercise, Instant Drill, Instant Drill Corner, Exercise, Supplementary Exercise, Class Activity, Inquiry & Investigation, and Unit Test. Within each category, there are usually many tasks. A *task* here means, following Stylianides (2009, p. 270), any problem in the exercises or parts thereof that have a separate marker. *Task*

Table 13.1 The analytic framework (Stylianides 2009, p. 262)

Reasoning-and-proving			
I. Making mathematical generalizations		II. Providing support to mathematical claims	
(a) Identifying a pattern	(b) Making a conjecture	(c) Providing a proof	(d) Providing a non-proof argument
1. Plausible pattern 2. Definite pattern	3. Conjecture	4. Generic example 5. Demonstration	6. Empirical argument 7. Rationale

served as unit of analysis in this study and there were totally 3241 tasks to be analyzed and categorized into the seven subcategories of the constituent RP activities set out in Table 13.1. In a manner similar to Hanna and de Bruyn (1999) and Otten et al. (2014a), to decide whether a task was an RP task, we looked at its form in the students' textbook (e.g., key phrases such as "Prove that…", "Explain your answer."). In cases where this was not clear, we consulted the Teacher's Manual (which contains suggested solutions, but only suggested solutions, to all the exercises) in order to infer what type of response was expected for students.

Examples of Analysis

Although there is a considerable amount of exercises in these algebraic chapters, as far as the coding process was concerned, tasks involving reasoning-and-proving opportunities (RPTs) were classified into three types: Type-1, Type-2 and Type-3.

(A) Type-1 RPTs explicitly ask for justification (or explanation) and their usual forms are "Prove that…" and "Explain your answer." (Philosophically, justifying may not be the same as explaining, see, e.g., Kasachkoff (1988), but we do not venture to go into this controversial issue here.)

Example 1 (see Fig. 13.1) belongs to Type-1 since it explicitly states "Explain your answer". While Task (a) is not a proof task, Task (b) is and is coded as "Demonstration—Proof by Calculation" as, according to the *Teacher's Manual 4B* (p. 15), the student has to provide a deduction (or calculation steps) as shown in Fig. 13.1 to justify (or explain) the answer. Reid and Knipping (2010, p. 124) call this proof method "mechanical deduction". Though this type of deduction involves mechanical algebraic manipulations and little reasoning, logically it should be regarded as proof (see also Slomson 1996, p. 11, "Proofs as Calculations").

(B) Type-2 RPTs implicitly ask for justification and their usual form is "Determine whether …" The aim of these RPTs is to check students' understanding of definitions.

Example 1

Let $f(x) = 2^x$ be a function.

(a) Find the values of $f(2)$, $f(3)$ and $f(6)$.
(b) Is the value of $f(6) \div f(2)$ equal to that of $f(3)$? Explain your answer.

Solution (from Teacher's Manual):

(a) $f(2) = 2^2 = 4$; $f(3) = 2^3 = 8$; $f(6) = 2^6 = 64$

(b) $f(6) \div f(2) = 64 \div 4 = 16$; $f(3) = 8$
$\because\ 16 \neq 8$
$\therefore$ the value of $f(6) \div f(2)$ is not equal to that of $f(3)$.

Fig. 13.1 Tasks 10(a) and 10(b) of Exercise 4B of Chapter 4 Basic Knowledge of Functions of Book 4A (p. 26)

Example 2 (see Fig. 13.2) belongs to Type-2 since it does not explicitly ask for justification. However, according to the *Teacher's Manual 4A* (p. 7), justification is required (by appealing to the definitions "zero is neither positive nor negative" and "any integer x can be written as $x/1$", respectively, for (a) and (b)). Both tasks are coded as "Demonstration—Proof by Definition". Task (c) can be proved to be false by giving a counter-example, e.g., "9", which is a real number but not irrational. This task is therefore coded as "Demonstration—Proof by Counter-Example".

(C) Type-3 RPTs are special in that they are usually templates for illustrating reasoning-and-proving. Tasks of this type are usually found in Class Activity or Inquiry & Investigation. In the process of coding, Type-3 RPTs are *dually* coded in that, on the one hand, the task is coded as a separate unit of analysis as in the coding of Type-1 and Type-2 RPTs, and, on the other hand, the task is coded as part of the constituent activity (or activities) of RP being illustrated.

Example 3 (see Fig. 13.3) belongs to Type-3, which means that it needs to be dually coded. First, each task is coded as a separate unit of analysis. In this example, none of the tasks 1(a), 1(b), 1(c) and 1(d) has any RP opportunities for the student. Task 2 is coded as "Conjecture". Then, each task is coded as part of the constituent activity (or activities) of RP being illustrated, which, in this example, is using an empirical argument to establish a definite pattern. So, tasks 1(a), 1(b), 1(c), 1(d) and 2 are all coded as "Empirical Argument" and "Definite Pattern".

Example 4 (see Fig. 13.4) also belongs to Type-3 and so it needs to be dually coded. First, each task is coded as a separate unit of analysis. In this example, none of the tasks 1, 2, 3, 4, 5, *conclusion* and *extension* has any RP opportunities for the student. Then, each task is coded as part of the constituent activity (or activities) of RP being illustrated, which, in this example, is using a generic example to prove a result. So, tasks 1, 2, 3, 4, 5 and *conclusion* are all coded as "Generic Example". Note that here the task *extension* is not part of the RP activity being illustrated and so is not coded as "Generic Example".

Example 5 (see Fig. 13.5) belongs to Type-3, which means that it needs to be dually coded. First, each task is coded as a separate unit of analysis. In this example, none of the tasks 1(a), 2(a), 2(b), 3, 4 and *conclusion* has any RP opportunities. Tasks 1(b) and 2(c) are coded as "Rationale" because there is a statement used in their respective arguments in the *Teacher's Manual 4A* (p. 26), which is not properly justified (namely, in 1(b), "a^2 is an even number and so a is

Example 2

Determine whether each of the following statements is true (T) or false (F).

(a) Zero is a positive integer.
(b) All integers are rational numbers.
(c) All real numbers are irrational numbers.

Fig. 13.2 Tasks 1(a), 1(b) and 1(c) of Exercise 1A of Chapter 1 Number Systems of Book 4A (p. 11)

Example 3

Objective: To investigate the relation among the dividend, the divisor, the quotient and the remainder in the division of polynomials.

1. Complete the following table.

	Dividend	*Divisor*	*Quotient*	*Remainder*	*Divisor × quotient + remainder*
(a)	$x + 2$	x			
(b)	$4x - 5$	$x + 1$			
(c)	$2x^2 - x + 1$	$x - 3$			
(d)	$x^3 + x^2 - x - 1$	$x + 2$			

2. From the results of Question 1, what do you think about the relation among the dividend, the divisor, the quotient and the remainder?

Fig. 13.3 Tasks of Class Activity 1 of Chapter 6 More about Polynomials of Book 4A (p. 11)

Example 4

Exploring the principle of the method of completing the square

Inquiry

Apart from comparison with the perfect square identity, is there any other way to explain the principle of the method of completing the square?

Investigation Steps

Try to make the quadratic polynomial $x^2 + 14x$ a perfect square according to the following steps.

1. Draw a square of side x units. What is its area?
2. Divide the coefficient 14 of x in the quadratic polynomial by 2 and get 14/2.
3. Add a rectangle of length x units and width 14/2 units to the top and the right sides of the square in Step 1 (see Fig. I). What is the total area of the figure?

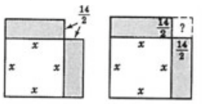


Fig. I   Fig. II

4. In order to make the figure in Step 3 form a larger square, what figure should be fitted in the top right corner of the figure (see Fig. II)? What is the area of this small figure?
5. What constant should be added to $x^2 + 14x$ in order to make it a perfect square?

Conclusion

For the quadratic polynomial $x^2 + bx$, area of the larger square + total area of the two rectangles + area of the smaller square $= x^2 + 2 \times$ _______ $\times x +$ (_______ $)^2$ = (________ $)^2$

Extension

Try to use the geometric method above to solve the quadratic equation $x^2 + 10x = 39$.

Fig. 13.4 Tasks of 5.1 Inquiry & Investigation of Chapter 5 Quadratic Functions of Book 4A (p. 44)

an even number too" and, in 2(c), "b^2 is an even number and so b is an even number too"). Task *extension* on its own is coded as "Rationale" since there is a statement, "a^2 is a multiple of 3 and so is a multiple of 3 too", which is not properly justified in its proof in the *Teacher's Manual 4A* (p. 26). Then, each task is coded as part of the constituent activity (or activities) of RP being illustrated, which is, in this example, using a proof by contradiction to show that 2 is an irrational number. So, tasks 1(a), *Example 5*

Exploring whether √2 is a rational number or an irrational number

Inquiry

Is √2 a rational number or an irrational number?

Investigation Steps

Let $\sqrt{2} = a/b$, where a and b are positive integers, and they do not have any common factor (except 1).

1. (a) By squaring both sides of $\sqrt{2} = a/b$, express a^2 in terms of b.
   (b) Using the result of (a), explain why 2 is a factor of a.

2. Since 2 is a factor of a, we can let $a = 2k$, where k is an integer.
   (a) By squaring both sides of $a = 2k$, express a^2 in terms of k.
   (b) Using the results of 1(a) and 2(a), express b^2 in terms of k.
   (c) Is 2 a factor of b? Explain your answer.

3. Using the results of 1(b) and 2(c), what is a common factor (except 1) of a and b?

4. Does the result in Question 3 contradict the assumption that a and b do not have any common factor (except 1)?

Conclusion

√2 is (a rational number / an irrational number).

Extension

Prove that √3 is an irrational number.

Fig. 13.5 Tasks of Inquiry & Investigation of Chapter 1 Number Systems of Book 4A (p. 32)

1(b), 2(a), 2(b), 2(c), 3, 4 and *conclusion* are all coded as "Demonstration—Proof by Contradiction". Task *extension* on its own is also coded as "Demonstration—Proof by Contradiction".

Results and Discussion

As shown in Table 13.2, there are relatively limited opportunities (410 out of 3241 tasks, i.e., 13%) as compared to, for example, 40% of Stylianides (2009) reviewed above, for students to learn reasoning-and-proving in the algebraic chapters in Year 10 of the chosen textbook series. This lack of attention to reasoning and proof in algebra is somewhat expected, given that the curriculum mentions proof only in the learning targets of geometry, namely, to "formulate and write geometric proofs involving 2-dimensional shapes with appropriate symbols, terminology and reasons" (Education Bureau HKSARG 2007, p. 15). Nonetheless, this lack of attention to reasoning and proof in algebra is still in contrast to the international call that reasoning and proof should permeate school mathematics at all levels and across all content areas in order for reasoning and proof to become a "habit of mind" (see, e.g., NCTM 2000, p. 56). A consequence that might be attributed to this lack of emphasis on reasoning and proof in the curriculum is that, as informed by TIMSS 2011 (Mullis et al. 2012, pp. 148 and 150), "Hong Kong students in general do well in Knowing items, and relatively badly in Reasoning items" (Leung 2015, p. 3).

Table 13.2 Frequency and distribution of RP tasks across RP subcategories

Reasoning-and-proving subcategory	Frequency (%)
I. Making mathematical generalizations:	46 (11.2)
(a) Identifying a pattern:	42 (10.2)
1. Plausible pattern	9 (2.2)
2. Definite pattern	33 (8.0)
(b) Making a conjecture:	4 (1.0)
3. Conjecture	4 (1.0)
II. Providing support to mathematical claims:	364 (88.8)
(c) Providing a proof:	304 (74.1)
4. Generic example	24 (5.9)
5. Demonstration	280 (68.3)
(d) Providing a non-proof argument:	60 (14.6)
6. Empirical argument	57 (13.9)
7. Rationale	3 (0.7)
Total	410 (100)

Furthermore, the majority of these reasoning-and-proving opportunities (364 out of 410, i.e., 88.8%) were categorized as Providing Support to Mathematical Claims, and only a small proportion (46 out of 410, i.e., 11.2%) were categorized as Making Mathematical Generalizations (in which there were only 4 instances of Conjecture). This shows not only that there were almost no opportunities for conjecturing in the algebraic chapters of the textbook series, but also that making mathematical generalizations (*identifying a pattern* and *conjecturing*) and providing support to mathematical claims (*proving*) were treated, in large part, in isolation from each other. This disjunction between making mathematical generalizations and providing support for those mathematical claims is not unique to the context of Hong Kong—see, for example, Stylianides (2009) and Davis et al. (2014) reviewed above for the case in the U.S., and Davis (this volume) for the case in Ireland. However, some would argue that such treatments are problematic, as these activities are both fundamental and interrelated aspects of doing mathematics (see, e.g., Boero et al. 2007; Cañadas et al. 2007; Pedemonte 2007). According to these research studies, some kind of continuity, called *cognitive unity*, exists between the construction of a conjecture and the construction of its proof in such a way that the argumentation activity developed to produce the conjecture can be used to construct the proof "by organizing in a logical chain some of the previously produced arguments" (Pedemonte 2007, pp. 24–25).

Furthermore, within the main category of Providing Support to Mathematical Claims, the majority of the reasoning-and-proving opportunities were Demonstration (280 out of 364, i.e., 77%). However, as shown in Table 13.3, out of these 280 demonstration proofs, 50% were Proof by Definition or Proof by Counter-Example like those used in our Example 2, whose aim is just to check students' understanding of definitions. Another 47% were Proof by Calculation,

Table 13.3 Frequency and distribution of proof methods used in Demonstration

Proof method	Frequency (%)
Proof by calculation	131 (47)
Proof by definition	111 (40)
Proof by counter-example	28 (10)
Proof by contradiction	10 (4)

that is, proof (or, explanation) by giving a deduction (or, calculation steps) without stepwise justifications as in our Example 1, though such deduction itself may not be explanatory—this point was also observed by Reid (1998) in the context of secondary school mathematics in China. All these demonstration proofs required little reasoning. On the other hand, those proofs and non-proofs requiring more substantive reasoning like Generic Example and Rationale were found only in a small number of instances of Class Activity or Inquiry & Investigation. If they were not covered in class due to time constraints, then the student would not have opportunities to learn these important types of reasoning. A final point is that there were many instances of Non-Proof—Empirical Argument (57), more than double those of Proof—Generic Example (24). This might mislead students into believing that an empirical argument is sufficient to establish truth in mathematics—a widespread misconception among students and even teachers (see, e.g., Stylianides and Stylianides 2009).

All of the above results seem to suggest that reasoning and proof plays only a marginal role in school mathematics in Hong Kong. Given that all those non-RP exercises in the textbook (which are the overwhelming majority) focused on practicing and memorizing mathematical concepts and procedures, the above results to some extent confirm the previous findings that in secondary school classrooms in Hong Kong students' activities mainly focus on practicing and memorizing mathematical concepts and procedures (e.g., Leung 2001). Some would argue that such a treatment of reasoning and proof in school mathematics is problematic, as it deprives students of opportunities to experience proof as a vehicle for sense making and understanding in mathematics education.

Interestingly, in spite of this, Hong Kong students still consistently outperform their Western counterparts in school mathematics—an instance of the so-called "Paradox of the Chinese learner" (Watkins and Biggs 1996). The fact that school mathematics textbooks in Hong Kong stress drilling of procedural skills to a great extent may be due to influences from Chinese culture (or, more specifically, the Confucian heritage culture or CHC) which believes that "the process of learning often starts with gaining competence in the procedure, and then through repeated practice, students gain understanding" (Leung 2006, p. 43) (see also Fan et al. 2004; Marton et al. 1996).

As a final remark, we want to point out that the RP opportunities we found in the textbook series are only potential opportunities, as the teacher may not assign all the exercises with RP opportunities to students, even though Hong Kong teachers rely heavily on textbooks in their teaching. Thus the results we obtained can only be claimed as a best-case analysis of the RP opportunities.

References

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

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

Boero, P., Garuti, R., & Lemut, E. (2007). Approaching theorems in grade VIII: Some mental processes underlying producing and proving conjectures, and condition suitable to enhance them. In P. Boero (Ed.), *Theorems in school: From history, epistemology and cognition to classroom practice* (pp. 249–264). Rotterdam, The Netherlands: Sense Publishers.

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

Cañadas, M. C., Deulofeu, J., Figueiras, L., Reid, D., & Yevdokimov, O. (2007). The conjecturing process: Perspectives in theory and implications in practice. *Journal of Teaching and Learning, 5*(1), 55–72.

Common Core State Standards Initiative (CCSSI). (2010). Common core state standards for mathematics. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers.

Davis, J. D. (2012). An examination of reasoning and proof opportunities in three differently organized secondary mathematics textbook units. *Mathematics Education Research Journal, 24,* 467–491.

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

De Villiers, M. (1990). The role and function of proof in mathematics. *Pythagoras, 24*(1), 17–24.

Education Bureau HKSARG. (2007). *Mathematics curriculum and assessment guide* (*Secondary 4–6*) (updated 2014). http://334.edb.hkedcity.net/doc/eng/curriculum/Math%20C&A%20Guide_updated_e.pdf.

Fan, L., Wong, N.-Y., Cai, J., & Li, S. (Eds.). (2004). *How Chinese learn mathematics: Perspectives from insiders* (Vol. 1). World Scientific.

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

Hanna, G., & de Bruyn, Y. (1999). Opportunity to learn proof in Ontario grade twelve mathematics texts. *Ontario Mathematics Gazette, 37*(4), 23–29.

Hanna, G., & de Villiers, M. (Eds.). (2012). *Proof and proving in mathematics education: The 19th ICMI study* (Vol. 15). Springer Science & Business Media.

Hunte, A. A. (2016). *Mathematics education reform in Trinidad and Tobago: The case of reasoning and proof in secondary school*. Doctoral dissertation, University of Illinois at Urbana-Champaign.

Kasachkoff, T. (1988). Explaining and justifying. *Informal Logic, 10*(1), 21–30.

Leung, F. K. S. (2001). In search of an East Asian identity in mathematics education. *Educational Studies in Mathematics, 47*(1), 35–51.

Leung, F. K. S. (2006). Mathematics education in East Asia and the West: Does culture matter? In F. K. S. Leung, K. D. Graf, & F. J. Lopez-Real (Eds.), *Mathematics education in different cultural traditions: A comparative study of East Asia and the West* (pp. 21–46). New York: Springer.

Leung, F. K. S. (2015). Keynote speech: What does TIMSS inform us about mathematics (under) achievers in Hong Kong? In K. L. Wong, et al. (Eds.), *Proceedings of the Hong Kong Mathematics Education Conference 2015*.

Leung, F., & Park, K. (2002). Competent students, competent teachers? *International Journal of Educational Research, 37*(2), 113–129.

Leung, F. K. S., Wong, C. S., Hung, F. Y., Wan, Y. H., Wong, T. W., Ding, W. L., et al. (2014–2016). *New century mathematics* (2nd ed.). Oxford University Press (China) Limited.

Marton, F., Dall'Alba, G., & Tse, L. K. (1996). Memorizing and understanding: The keys to the paradox. In *The Chinese learner: Cultural, psychological and contextual influences* (pp. 69–83).

Mullis, I. V., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 International results in mathematics*. Amsterdam: International Association for the Evaluation of Educational Achievement.

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

Nordström, K., & Löfwall, C. (2005). Proof in Swedish upper secondary school mathematics textbooks—The issues of transparency. In *Proceedings of the 4th Congress of European Research in Mathematics Education* (*CERME4*) (pp. 448–457).

Otten, S., Gilbertson, N. J., Males, L. M., & Clark, D. L. (2014a). The mathematical nature of reasoning-and-proving opportunities in geometry textbooks. *Mathematical Thinking and Learning, 16*(1), 51–79.

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

Pedemonte, B. (2007). How can the relationship between argumentation and proof be analyzed? *Educational Studies in Mathematics, 66*(1), 23–41.

Reid, D. A. (1998). Sharing ideas about teaching proving. *Mathematics Teacher, 91*(8), 704–707.

Reid, D. A., & Knipping, C. (2010). *Proof in mathematics education: Research, learning and teaching*. Rotterdam, The Netherlands: Sense Publishers.

Slomson, A. (1996). Mathematical proof and its role in the classroom. *Mathematics Teaching, 155,* 10–13.

Stacey, K., & Vincent, J. (2009). Modes of reasoning in explanations in Australian eighth-grade mathematics textbooks. *Educational Studies in Mathematics, 72*(3), 271–288.

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

Stylianides, G. J. (2009). Reasoning-and-proving in school mathematics textbooks. *Mathematical Thinking and Learning, 11*(4), 258–288.

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

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

Tam, H. P., Wong, N. Y., Lam, C. C., Ma, Y., Lu, L., & Lu, Y. J. (2014). Decision making in the mathematics curricula among the Chinese mainland, Hong Kong, and Taiwan. In *Mathematics curriculum in school education* (pp. 93–117). The Netherlands: Springer.

Thompson, D. R., Senk, S. L., & Johnson, G. J. (2012). Opportunities to learn reasoning and proof in high school mathematics textbooks. *Journal for Research in Mathematics Education, 43*(3), 253–295.

Watkins, D. A., & Biggs, J. B. (Eds.). (1996). *The Chinese learner: Cultural, psychological, and contextual influences*. Hong Kong: Comparative Education Research Centre.

Wong, K.-C. (2017a). *Reasoning-and-proving in geometry in school mathematics textbooks in Hong Kong. CERME10*, Dublin, February 1–5, 2017.

Wong, K.-C. (2017b). *Reasoning and proof in probability and statistics in school mathematics textbooks in Hong Kong. PME41*, Singapore, July 17–22, 2017.

Chapter 14
Irish Teachers' Perceptions of Reasoning-and-Proving Amidst a National Educational Reform

Jon D. Davis

Abstract The syllabi driving the secondary mathematics education reform in Ireland expect students to engage in two components of reasoning-and-proving (RP) (Stylianides in For Learn Math 28:9–16, 2008): making mathematical generalizations (pattern identification and conjecturing) and providing support to mathematical claims (providing a proof/non-proof argument). This study examines the perceptions of pattern identification, conjecturing, and proof by 22 Irish teachers with varying levels of teaching experience via semi-structured interviews. These teachers perceived pattern identification and conjecturing as disconnected from proof construction. Indeed, teachers struggled to define conjecturing and proof. There also appeared to be a bifurcation in students' classroom experiences with RP processes. Teachers stated that the experiences with proof of students with perceived lower ability levels ended at pattern identification while higher-level students rarely engaged in pattern identification and focused on memorizing proofs due to the influence of high stakes assessments. The implications of these results are discussed.

Keywords Proof-and-reasoning · Curriculum · Reform · Teacher perceptions

Introduction and Background

Given that proof has historically appeared solely in secondary school geometry (Fujita and Jones 2014; Herbst 2002), teachers tend to consider proof as residing within this content area (Furinghetti and Morselli 2011). An expansive body of research has documented the struggles that students experience with proof (e.g., Healy and Hoyles 2000) and the deleterious effects of traditional proof instruction in geometry (Schoenfeld 1989). Lampert's (1990) work with fifth-grade students illustrated that it is possible for the teacher and students to co-construct a classroom environment where students make conjectures and engage in the development of

J. D. Davis (✉)
Western Michigan University, Kalamazoo, USA
e-mail: jon.davis@wmich.edu

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_14

valid mathematical arguments. Despite such proof of concept studies, teachers encounter similar struggles in identifying valid proofs (Knuth 2002a) and understanding the limits of empirical evidence (Martin and Harel 1989). Teachers also hold conceptions about the teaching of proof. For instance, in addition to formal proof, which was suitable only for some students, teachers also believed that proofs could be informal, consisting of empirical arguments (Knuth 2002b), which were suitable for all students. Indeed, teachers may harbor the conception that only students of perceived high ability can construct proofs (Furinghetti and Morselli 2011). Additionally, Dickerson and Doerr (2014) found that a group of 17 high school mathematics teachers stated that the most important proof purposes were: enhancing students' understanding of mathematics and developing generalizing skills that could be used in other areas. Knuth (2002a) found that a group of 16 practicing secondary school mathematics teachers identified proof as having a number of different purposes (e.g., establishment of truth), but did not see how proof could promote understanding. Knuth (2002b) found that the majority of teachers in his study defined proof as a logical or deductive argument demonstrating the truth of an assertion. The remaining teachers defined proof as a convincing argument.

Although there are a number of different ways to define understanding, one way is via relational and instrumental understanding (Skemp 1976). Skemp defined relational understanding as knowing both what to do in particular situations as well as why it is important to do those things. Instrumental understanding, in contrast, involves knowing rules without the knowing the reasons behind these rules. DeVilliers' (1991) work with 17 14-year-old pupils' exploration of geometric ideas in a dynamic geometry system found that while they were convinced of the validity of the ideas after a few manipulations, they did desire to understand why these ideas held true. Thus, even when convinced that a mathematical idea is true, students are still interested in gaining a relational understanding of these ideas. Such findings suggest that students' desire for relational understanding can be leveraged to motivate them to engage in proofs that serve an explanatory role.

In 2010 Ireland implemented Project Maths, a reform of their secondary mathematics educational system intended to promote student learning through problem-solving, investigations, and the use of real-world contexts. Research involving teachers who piloted this reform suggest that it involves practices that are a significant departure from previous classroom instruction in Ireland (National Council for Curriculum and Assessment [NCCA] 2012). For instance, teachers spoke about rote learning, memorization, a focus on mathematical procedures, and a presentation of procedures without justification before implementation of Project Maths. Two representations of the Project Maths reform are the syllabi for junior certificate students (ages 12–15) and senior certificate students (ages 16–18), each of which contains content objectives separated by three ability levels listed in ascending order of difficulty: foundation level (FL), ordinary level (OL), and higher level (HL). The standards appearing in these syllabi were designed for teachers to use as they craft classroom instruction. Reasoning-and-proving (RP) processes (Stylianides 2008) involving identifying patterns (plausible and definite),

Students learn about	Students should be able to
1.8 Synthesis and problem-solving skills	– explore patterns and formulate conjectures – explain findings – justify conclusions – communicate mathematics verbally and in written form – apply their knowledge and skills to solve problems in familiar and unfamiliar contexts – analyse information presented verbally and translate it into mathematical form – devise, select and use appropriate mathematical models, formulae or techniques to process information and to draw relevant conclusions.

Fig. 14.1 Synthesis and problem-solving skills from the syllabus for senior certificate students (NCCA 2014)

constructing conjectures, and providing support for mathematical claims appear throughout these documents (Davis 2014), but are especially singled out in a section titled, Synthesis and Problem-Solving Skills, that appears in each content strand as seen in Fig. 14.1. Figure 14.1 reveals that students are expected to identify patterns and develop conjectures. Students are asked to craft arguments when they explain their findings, justify their conclusions, and draw relevant conclusions.

In research reported elsewhere (Davis 2014) I found that junior certificate students were provided with fewer opportunities to develop conjectures and develop arguments than senior certificate students. In addition, each syllabus contained a preponderance of pattern identification, conjecture formulation, and proof construction outcomes that were not tied to specific mathematics content objectives. These standards were designed for teachers to use as they craft classroom instruction and given these disparities teachers may provide students with fewer opportunities to engage in RP during the enacted curriculum. The presence of RP processes in all content strands in this section of the syllabus led to questions surrounding teachers' conception of the role of proof in different content strands.

The centrality of RP processes in the development of mathematical knowledge by mathematicians has been pointed out by Stylianides and Silver (2009) in citing the work of Polya (1954), Atiyah (1984/1986) and Schoenfeld (1983). Research by Stylianides (2008) on one set of middle school textbooks found that 10% of tasks involved pattern identification, 1% of tasks involved conjecturing, and 5% of tasks involved demonstrative arguments. Otten et al. (2013) examined six US geometry textbooks for the presence of RP processes and found that the majority of these elements appeared on introductory proof units with less than 5% of exercises in other chapters involving RP processes. Wong and Sutherland (this volume) also found that 13% of tasks in a popular Hong Kong grade 10 algebra textbook series involved RP processes. Research involving US students' experiences with proof in geometry suggests that teachers skip tasks involving proof (Thompson and Senk 2014), rarely use technology such as dynamic geometry (Thompson and Senk), or reduce the cognitive demand of problems requiring arguments when enacted (Sears and Chávez 2014). Thus, given these findings it is likely that school students do not routinely experience pattern identification, conjecturing, and proof as a coordinate set of activities. Moreover, it is important to note that these RP processes do not necessarily have to follow a linear sequence. Karunakaran (this volume) has pointed out that mathematicians may interrupt their deductive logic in their process in order to complete their deductive arguments.

This study provides a unique contribution to the research body by using Stylianides' (2008) framework to examine a group of Irish teachers' mathematical, psychological, and pedagogical perspectives about RP processes and the connections between pattern identification and proof and the connections between conjecturing and proof. This study also sheds light on how mathematics teachers in transition from one set of educational experiences to another perceive RP processes. Specifically, this study was designed to answer the following research questions:

1. How does a group of Irish post-primary teachers perceive pattern identification and conjecturing and their roles in the development of proof arguments?
2. How does a group of Irish post-primary teachers define proof, position proof with regard to a mathematics content area, and envision the role of proof for students with different perceived abilities?
3. How does a group of Irish post-primary teachers conceive the role of proof in promoting understanding?
4. How does a group of Irish post-primary teachers perceive of the synthesis and problem-solving skills section of the syllabus that contains RP processes?

Methods

Stylianides (2008) introduced the term reasoning-and-proving (RP) to encompass the interconnected nature of four activities: pattern identification; conjecturing; constructing non-proof arguments; and developing proof arguments. These processes involve three components: mathematical, psychological, and pedagogical. The *mathematical component* consists of the examination of an activity for the presence of the aforementioned RP processes by a mathematically knowledgeable person. The *psychological component* seeks to understand how an individual who is engaged in solving a mathematical problem perceives RP processes. The *pedagogical component* has two subcomponents. First, it seeks to discern the relationship between the mathematical component of a problem with the solver's psychological component of a problem. Second, it seeks to understand the teacher moves that will enable the mathematical component of a task to become transparent to the solver.

This study makes use of all three components within Stylianides' (2008) framework. The focus of the psychological component is the teacher as solver and in the pedagogical component the teacher assumes the role of knowledgeable other with the interviewer seeking to understand the teacher moves that he or she takes to help students understand the mathematical nature of tasks involving RP processes. Participants were not presented with particular tasks, but in the course of the semi-structured interviews (described in more detail below) teachers either presented mathematical tasks as examples on their own or were pressed to provide examples. These tasks were then examined from a mathematical perspective to better understand the validity of the psychological component exhibited by the teachers and the nature of the teacher moves taken to make the mathematical perspective transparent to students (pedagogical component). If teachers did not provide examples and instead talked about RP processes in general, these descriptions were compared to their descriptions in Stylianides. This framework was chosen as the Project Maths syllabi contain these different RP processes (Davis 2014).

Post-primary teachers working with senior cycle mathematics students (ages 15–18) within a 50-mile radius of Limerick, Ireland were recruited for this study. The teaching experience and gender of these teachers appear in Table 14.1. teaching experience was broken down into three categories: low (0–5 years), medium (6–14 years), and high (15–30 years). Interview questions associated with each of the research questions appear in the Appendix. The semi-structured nature of the interviews enabled the author to follow-up on some teachers' responses when they lacked sufficient details. Participants would occasionally provide examples from their classes where they engaged their students in RP processes. Twenty-two teachers were asked about pattern identification, 18 teachers were asked about conjecturing, and 21 teachers were asked about proof during semi-structured interviews lasting between 30 min and one hour during the 2012–2013 school year. The interview length varied as teachers were interviewed during the school day and had variable amounts of time for the interview. The variable interview lengths as well as the semi-structured nature of the interviews led to differing numbers of teachers answering the different questions associated with the interview questions. Each interview was audiotaped and transcribed and coded using HyperResearch (Researchware 2011) qualitative analysis software. Qualitative data analysis methods of analytic induction and constant comparison were used to identify and refute interpretations for themes and relationships appearing within the data (Miles et al. 2014).

Table 14.1 Participating teachers' experience and gender

Experience		
Low	Medium	High
6	8	8

Gender	
Female	Male
12	10

Results

Pattern Identification

Seventeen out of twenty-two teachers most closely connected pattern to a specific mathematics unit, Sequences and Series, as seen in the following interview excerpt.

> Interviewer: Is the identification of patterns connected to the construction of proofs at all?
> Aidan: Um, I'm being brutally honest but, I haven't taught the patterns, sequences, series [sic] section yet.

This excerpt also illustrates an important problem with connecting patterns with a specific unit, namely, students may experience a great deal of instruction before engaging in pattern identification. It also promotes in teachers a perspective that patterns are not a unifying mathematical activity that occurs across a variety of mathematics content areas. This stands in contrast to the Synthesis and Problem-Solving Skills section that appears in each content strand of the syllabus. The particular unit mentioned by the teacher in this excerpt provides students with opportunities to identify patterns in arithmetic and geometric sequences. Indeed, in the senior cycle ordinary level teacher handbook for Project Maths this is the only unit out of twelve that contains the word "patterns" in its title.

Additionally, sixteen of the teachers did not see pattern identification as a step leading towards proof. Indeed, for eight of these teachers the identification of plausible patterns was synonymous with proof, especially for students learning mathematics at the FL and OL.

> Cassidy: Ah, another example that I can think of is okay when I'm doing the laws of indices. You know and you work away with them and then you come to this one and any number to the power of 0 is equal to 1. And, I kind of turn around to them and I say well, I mean do you not think that looks a bit crazy? Where are they getting that from you know? Anything to the power of 0 equals 1. So I go back and I'd say, "Well actually I'll show you where it comes from. You know you don't need to learn this or anything off by heart but just it's interesting to look at." And maybe I'll [do] 2 cubed and 2 squared and 2 to the power of 1. And then I'll get them to type into their calculators, uhm 2 to the power of minus 1, 2 to the power of minus 2, 2 to the power of minus 3 and we'll take a look at the kind of pattern that emerges. And if you work it from either side, from let's say the positive powers to the negative powers you can easily see by the pattern when you have 2 to the power of 0 what you should get is 1, uhm, as your answer.
> Interviewer: So does that prove that any number to the 0 power is 1?
> Cassidy: Well, you can try other numbers. And you keeping getting the same answers so we have done it, we'll do 2 and I'll say, "Okay we'll check out to see if this works for three and then it does prove that any number to the power of 0 equals 1." I think anyway. I think it proves it enough to them that they can accept it then.
> Interviewer: Okay, proves that it is always true.
> Cassidy: Yes.
> Interviewer: You use the words "enough for them" would it be different for you?
> Cassidy: No. Personally, I would accept that.

Indeed, the excerpt above suggests that Cassidy was not likely to see pattern identification as a necessary step in moving towards a proof, as pattern identification appeared to be sufficient for her in showing that a mathematical idea always held true. According to the teachers, HL students, on the other hand, were more likely to spend less time with pattern identification and focus more on the construction of proofs. In the excerpt below, Ashling confirms the different types of RP activities in her classroom for students of different perceived abilities. In this excerpt, "activity" refers to an opportunity for students to identify a mathematical pattern.

> Interviewer: So with the ordinary level classroom you'd focus more on the activity?
> Ashling: Yes.
> Interviewer: So, given a finite amount of time in the classroom, you focus more on the activity and then less time on the formal proof?
> Ashling: Yes.
> Interviewer: Okay.
> Ashling: But in the ordinary group [sic].
> Interviewer: Okay, with the ordinary level. And maybe with the higher level it's less time on the activity and more time on the proof?
> Ashling: Yes.

Conjecturing

Out of 18 teachers asked about conjecturing, nine teachers were not able to define it at all. Four teachers provided incorrect definitions such as checking the reasonableness of an answer, assumptions when beginning a proof (2 teachers), and corollary. Two teachers were able to provide correct definitions of conjecture. Three teachers' responses were coded as unknown as it was not possible to categorize their definitions. Due to time constraints during the interview I only provided five of the nine teachers with the definition of a conjecture as an educated guess. Even when provided with the definition these teachers did not connect conjecturing with the object of proof development. Consider the excerpt below.

> Interviewer: Some people define a conjecture as a reasonable guess, or a reasonable hypothesis of what might happen.
> Seamus: An educated guess.
> Interviewer: Yeah, more or less. So does that play a role in the construction of a proof?
> Seamus: It does, yeah.
> Interviewer: Can you give an example of where students have done that in your classroom?
> Seamus: Well they can rule something out. I know we were talking about the different methods of approaching a question. Well you could use conjecture quickly to rule something out whether it will work or not to solve the proof.

In this excerpt, the teacher considers conjecturing as a way to rule out a method to construct a proof. One of the four teachers who incorrectly defined conjecturing also connected this act to problem-solving. When cast in this light conjecturing was seen as a check on the reasonableness of the answer or how to approach the solving of a problem. The other three teachers connected conjecturing to proof, but in nonstandard ways like the following: valid assumptions that can be made at the beginning of a proof; confusing conjecture with corollary; and confusing conjecture with axiom. None of the 18 teachers asked about conjecturing saw a connection to proof in terms of the object of a proof. The excerpt below is an example of this phenomenon.

Interviewer: So I hear you talking about conjectures in terms of making sure that the answer that you get from the calculator is reasonable, so going back to the example from geometry where the sum of the two interior angles equals the exterior angle, you show them a lot of examples in an ordinary level class. Do you then ask them to make a guess about whether they think it's always true or do you just go right to the proof then?
Simon: Personally I would say that I ordinarily go right to the proof.

Proof Development

Eleven out of 21 teachers defined proof as a logical derivation similar to what was found by Knuth (2002b). This is seen in the following excerpts from Ashling and Murtagh.

Interviewer: Our study is really focusing on reasoning-and-proving so something we have been asking teachers is your definition of what a mathematical proof is. So how would you define that?
Ashling: A set of logical steps and come out with an answer that's true.
Interviewer: So the focus of our study is really on reasoning-and-proving. And so, what we've been asking the teachers for is their definition of a mathematical proof.
Murtagh: I suppose a proof involves a number of logical steps that you follow to reach a conclusion. There might be some prerequisite knowledge in those steps or there might be some previous theorems or proofs that you already understand, or there might be an axiom that doesn't need to be proved, but you take those pieces of information and you piece them together until you can reach a logical conclusion about what you set out to prove to begin with.

Five teachers were not able to define a proof, but were only able to describe proof techniques such as proof by contradiction or provided theorems as definitions of proofs such as the irrationality of two. An example of the latter is seen below.

Interviewer: …How would you define what a mathematical proof is?
Shauna: Mathematical proof?
Interviewer: Ah huh.
Shauna: Well, something that has been well, discovered, years ago that um, say for example like um, like there's, the angles of a triangle add up to 180. That can be proven pretty easily and it doesn't necessarily have to be just taken as a given. You know you can take up all the triangle [sic] into the three angles and see that yourself or measure them with a protractor. You know in junior ones they'd cut off all the angles at the three corners and then actually put them together, make them 180. Things like that, but also then there's lots of ones then

that are taken, that's given really without proving it like say the area of a triangle is half b times c. It would be a lot more complicated to prove that, but it can be done.

Shauna's response also implies that a proof can be constructed through a series of examples such as cutting the corners off a set of triangles and showing that they form 180°. Four additional teachers considered a set of examples to denote a proof and one teacher's definition was connected to a pattern and especially how that pattern could be converted into a proof. If teachers did not provide specific examples of proof within a content area, they were asked which content areas proof appears in. These teachers were then asked a follow-up question about other content areas where proof appears. If teachers embedded their proof examples in a specific content area then teachers were asked about other content areas where proof can occur in school mathematics. The most popular content area mentioned by 20 out of 22 teachers in which proof occurred was geometry.

Kaitlin: Well, it's again I think proof, I think Geometry, very much Geometry, I think theorems I mean that's what I think when I think proof. Do you know? But other than that I don't really.

The next most frequent category was number (6 out of 22 teachers), followed by trigonometry (4 out of 22 teachers), and algebra (2 out of 22 teachers). A total of 18 teachers were asked a follow-up question about proof in other areas. Of these teachers only four mentioned algebra or provided valid examples of proof in algebra; four teachers mentioned calculus such as the use of first principles to prove differentiation formulas. In order to further probe teachers' understanding of proof in other content areas a total of ten teachers were asked specifically about proof in the statistics and probability content strand. One teacher replied that she did not know if proof appeared in this content strand. Two teachers provided invalid examples of proof in statistics and probability, three teachers did not answer the question, and three teachers stated that proof did not appear in this content strand. This suggests that teachers either do not construct proofs for probability ideas such as the addition rule or, if they do justify these ideas, they do not consider them to be proofs. That is, their definitions for proof while devoid explicitly of a content area as in the logical derivation definition above may indeed be context dependent.

A total of 15 teachers were either asked about proof for students with different perceived ability levels or described proof experiences for students of different perceived ability levels. Fourteen teachers stated that HL students had very different proof experiences than OL and FL students. For instance, Catherine noted that she would use the word "proof" among HL students and not with FL students because the use of this terminology would inspire anxiety in FL students. Hugh stated that HL students worked on more difficult proofs and four teachers noted that these proofs would be more abstract than those worked on by OL or FL students. Kerry not only noted that HL students would be asked to create proofs, but that they also would be expected to create diagrams to assist in developing proofs. Additionally, Kian noted, that HL students would be assessed on their ability to prove on the high stakes exams given at the end of the senior cycle. However, due to the presence of proof on these exams for HL students Fiona admitted that she would delay the instruction of proof until right before the exams. Brian and Iona admitted that because HL students would be asked to reproduce proofs that they had already seen the students would resort to memorizing these proofs with less of an emphasis on understanding. Aidan noted that neither FL nor OL students were asked to construct proofs in his classroom. As four teachers stated, OL students would most likely encounter proof through teacher led presentations. Three teachers stated that FL students' experiences with proof centered around applying the results of a theorem to find an answer such as a missing angle.

Connection Between Proof and Understanding

In order to ascertain teachers' conceptions about proof a total of twelve teachers were asked if proof promotes understanding. Two teachers felt that proof did not promote students' understanding of mathematical ideas. Hugh's definition of understanding focused on applying mathematical ideas. He elaborated that as the development of a proof did not involve its application, proof was disconnected from understanding. Four teachers felt that proof was tied to understanding, but that understanding was compartmentalized to developing a proof. That is, a proof required understanding on the part of students, but that understanding did not necessarily apply to other areas of mathematics. Two other teachers stated that proof was connected to understanding the origin of mathematical objects. Two other teachers noted that the results of proof were used in solving problems oftentimes set within real-world contexts and therefore promoted understanding. One teacher stated that proof involved verification of mathematical ideas (e.g., knowing that), but not necessarily knowing why mathematical ideas were true. This delineation is similar to Skemp's (1976) descriptions of instrumental and relational thinking. The last teacher believed that proof promoted understanding due to the fact that constructing a proof required making connections. This work enabled students to make connections across different mathematical ideas, which for this teacher was emblematic of understanding.

Connection to Syllabus Vis-à-Vis Synthesis and Problem-Solving

Recall that the *Synthesis and Problem-Solving* section of the syllabus contained a number of reasoning-and-proving components. A total of 22 teachers were asked what the section of the syllabus titled *Synthesis and Problem-Solving* meant to them in order to understand teachers' perceptions of this section of the syllabus. Twenty of the teachers did not connect this section to RP. Nearly all of the teachers did not recognize this section as connected to the syllabus. Instead they parsed this title into its constituents: problem-solving and synthesis. They described problem-solving as best they could, but when it came to synthesis eight teachers did not know how to define this terminology. Two teachers did connect problem-solving and synthesis to a component of RP, proof. These connections are seen in the following two excerpts.

> Nora: Synthesis, I'm not sure about that word, but um, problem-solving I suppose, trying to work through some logical steps to get from start to the end of the problem, I suppose.
> Kerry: Ah, I suppose kind of understanding the question and reading the words and, um, creating if you like, like I said, an equation or a table and making sense of or, the synthesis part is kind of the creating the table or drawing the graph, and then they're analyzing, and they're arriving at their answer, or they're justifying their answer saying why it's true or not true. So that would be my understanding of the synthesis and the problem-solving.

Discussion

Teachers' Perceptions of RP and Students' Abilities

The Project Maths leaving certificate syllabus distinguishes different content goals for students at the FL, OL, and HL levels. However, the syllabus makes no distinctions when it comes to students' experiences with three key RP processes: pattern identification, conjecture development, and the development of arguments. It is expected that all of these students be given opportunities to engage in all three of these processes. Yet, the interviews suggested that there is a bifurcation in classroom experiences for FL/OL and HL students, neither of which supports the goals of the syllabus nor promotes the interconnected nature of RP processes. FL and OL students were less likely to engage in the construction of arguments and HL students were less likely to identify patterns due to the influence of high stakes exams. FL/OL students' experiences in RP typically involved pattern identification. A sizeable minority of teachers felt that this pattern identification was sufficient for these students to show that mathematical ideas always held true and that pattern identification promoted student understanding. Yet the examples given of patterns suggested that they showed that something was true but not necessarily why it was true. The presence of pattern identification as a form of proof is similar to an empirical proof scheme as noted in previous research (Knuth 2002b; Martin and Harel 1989). While HL students were more likely to engage in the construction of proofs, these arguments were likely to be memorized by students and taught in a way to promote their quick recall due to high stakes assessments. This memorization is akin to the instrumental understanding (Skemp 1976) displayed by sixth grade students in Askevold and Lekaus (this volume). These results are similar to the beliefs held by a group of grade 10 geometry students in the United States with respect to proof (Schoenfeld 1989). The influence of these high stakes examinations could be harnessed by Project Math personnel to more closely promote the goals of the leaving certificate syllabus by including questions where FL, OL, and HL engage in pattern identification, conjecturing, and proof development across of range of content areas.

Teachers' Perceptions of RP as an Interconnected Set of Processes

The results of this study suggest that teachers do not view RP as a set of interconnected actions. Pattern identification was not connected to conjecturing or proof and for some teachers, pattern was synonymous with proof suggesting that for these teachers empirical arguments appeared to constitute valid proofs as researchers have found at the high school level in the USA (Knuth 2002b). Conjecturing was not connected to pattern identification or to proof. Indeed, teachers struggled to define conjecture and when provided with a definition they were more likely to connect it to problem-solving than proof or pattern identification. Teachers' disconnections are further supported by their responses that pattern occurred primarily in one mathematical unit and proof appeared predominantly in geometry. The latter finding is similar to what has been reported in the body of research around proof (e.g., Herbst 2002). This fixation on proof in geometry also appears in curriculum from other places around the world such as Hong Kong (Wong and Sutherland, this volume). These findings as well as teachers' struggles to describe the synthesis and problem-solving section appearing within each content strand of the syllabus suggest that teachers are not attending to this section of the leaving certificate syllabus. While Stylianides (2008) saw RP processes as interconnected and part and parcel to the work of mathematicians, the teachers in this study did not. This may reflect how these teachers themselves experienced instruction as students in secondary mathematics classrooms in Ireland or in the university where instruction may focus more on finished proofs and not how those proofs were developed (Stylianou et al. 2015). Consequently, professional development for Irish post-primary teachers could be fashioned around these processes to assist these individuals in understanding the synthesis and problem-solving skills that appear in each content strand, point out the ubiquitous of these processes, and provide examples to teachers for how they can implement these pattern identification, conjecturing, and developing arguments in the secondary mathematics classroom across a variety of different content strands.

Teachers' Perceptions of Proof

The teachers interviewed for this study held a number of perceptions about proof. The majority of teachers defined proof as a logical derivation similar to teachers from Knuth's (2002b) study had found. This study provides more evidence of teachers' struggle with proof. That is, teachers struggled in defining a proof, considered proof techniques such as proof by contradiction to be a definition of proof, or confused a proof with the products of a proof (e.g., theorems). Five teachers saw examples as valid proof, which resonates with previous research studies (e.g., Knuth 2002b). Additionally, the teachers in this study tended to connect proof with geometry similar to previous research (Furinghetti and Morselli 2011). Taken as a group these findings suggest that teachers may be fixated on the instructional aspects of proof. That is, they may see proof primarily through their work with students and not as learners of mathematics or as individuals who engage in mathematical work and make use of proof in that work. This is seen in their definitions of proof as methods to complete a proof (e.g., proof by contradiction) or the use of examples to illustrate the result of a proof that they use in their daily work with students.

This work builds on a previous study completed by Knuth (2002a) and Dickerson and Doerr (2014) as it seeks to describe teachers' perceptions with regard to the connections between proof development and understanding and it also seeks to further explicate teachers' conceptions with regard to a lack of connection between proof and understanding. Knuth found that a group of sixteen teachers did not view understanding as a role that proof can take in mathematics. However, the vast majority (16/17) of teachers interviewed by Dickerson and Doerr saw proof as promoting understanding in a number of different ways. This study found the number of teachers making connections between proof and understanding to lie between these two studies. This study found that five teachers did not connect proof construction to understanding. Four of these teachers saw understanding as connected to the use of mathematics to solve problems set within real-world contexts, so called applications of mathematics. Recall that this was one of the goals of the Project Maths reform.

Seven teachers saw proof as connected to understanding. That understanding was constrained to proof construction for several teachers, but three teachers saw proof as promoting understanding in a number of interesting ways such as the origin of mathematical ideas and connections across mathematical ideas. In addition, teachers also saw proof as promoting students problem-solving capabilities. These findings suggest other ways that teachers perceive the connections between proof and understanding besides those described by Dickerson and Doerr (2014) such as proofs developing transferrable thinking skills or metacognitive thinking skills. Additionally, these findings suggest that while teachers may harbor unproductive conceptions around proof such as seeing it as limited to geometry, they can also possess progressive conceptions such as linking proof with understanding. Such connections to understanding can be leveraged by those wishing to engineer professional development to help teachers see the value of engaging all students in the development of support for mathematical claims.

Conclusion

This study examined the mathematical, psychological, and pedagogical components of RP processes as seen through the eyes of a group of Irish post-primary teachers. Although mathematicians invoke these processes as they develop mathematical ideas, teachers themselves did not see these processes as connected at all. This study serves as another testament to the disjuncture between school mathematics and mathematics as practiced by mathematicians (Schoenfeld 1989). While this study adds to our understanding of teachers' perceptions about RP processes and their interconnectedness, it also speaks more broadly to curriculum reform and teachers' perceptions. Although RP processes appear in the leaving certificate syllabus, teachers were not cognizant of this component of the syllabus as it did not contain specific mathematics content. This finding serves as a strong argument in favor of interweaving valued mathematical processes with content standards in the design of national curriculum documents.

Appendix

Research question	Interview questions
(1) How does a group of Irish post-primary teachers perceive pattern identification and conjecturing and their roles in the development of proof arguments?	• Does the identification of patterns ever play a role in constructing a proof? • What is your definition of a conjecture? • Do you think conjectures have any role in proof as you define it and as you understand it?
(2) How does a group of Irish post-primary teachers define proof, position proof with regard to a mathematics content area, and envision the role of proof for students with different perceived abilities?	• What is your definition of a mathematical proof? • You have given me an example of proof in the area of geometry and it's linked to learning mathematics, understanding mathematics. Does proof play as important a role in learning other content areas as it does in geometry? • Is your definition of a proof the same for students in foundation, ordinary, and higher levels?
(3) How does a group of Irish post-primary teachers conceive the role of proof in promoting understanding?	• Does proof play a role in understanding as it is described in Project Maths? If so, how? If not, why not?
(4) How does a group of Irish post-primary teachers perceive of the synthesis and problem-solving skills section of the syllabus that contains RP processes?	• Something that appears in the leaving certificate syllabus at the end of each of the content strands is something that is called synthesis and problem-solving. Can you describe what that means to you?

References

Atiyah, M. F. (1984/1986). Mathematics and the computer revolution. In A. G. Howson & J. P. Kahane (Eds.), *The influence of computers and informatics on mathematics and its teaching*. Cambridge: Cambridge University Press.

Davis, J. D. (2014). Reasoning-and-proving within Ireland's reform-oriented national syllabi. *The Mathematics Enthusiast, 11*(3), 665–706.

DeVilliers, M. (1991). Pupils' needs for conviction and explanation within the context of geometry. In *Proceedings of the 15th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 255–262). Assisi, Italy.

Dickerson, D. S., & Doerr, H. M. (2014). High school mathematics teachers' perspectives on the purposes of mathematical proof in school mathematics. *Mathematics Education Research Journal, 26,* 711–733.

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

Furinghetti, F., & Morselli, F. (2011). Beliefs and beyond: Hows and why in the teaching of proof. *ZDM Mathematics Education, 43,* 587–599.

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education, 31*(4), 396–428.

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

Knuth, E. (2002a). Secondary school mathematics teachers' conceptions of proof. *Journal for Research in Mathematics Education, 33*(5), 379–405.

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

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal, 27*(1), 29–63.

Martin, W. G., & Harel, G. (1989). Proof frames of preservice elementary teachers. *Journal for Research in Mathematics Education, 20*(1), 41–51.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Thousand Oaks, CA: Sage.

National Council for Curriculum and Assessment. (2012). *Project maths: Reviewing the project in the initial group of 24 schools—Report on school visits.* http://www.ncca.ie/en/Publications/Reports/Project_Maths_Reviewing_the_project_in_the_initial_group_of_schools.pdf. Accessed: March 20, 2017.

National Council for Curriculum and Assessment. (2014). *Leaving certificate syllabus*. https://curriculumonline.ie/getmedia/fd79ce76-9a07-42fb-9dd7-ac6036861816/SCSEC25_Maths_syllabus_examination_in_2014_eng.pdf. Accessed: February 10, 2017.

Otten, S., Males, L. M., & Gilbertson, N. J. (2013). The introduction of proof in secondary geometry textbooks. *International Journal of Educational Research, 64,* 107–118.

Polya, G. (1954). *Mathematics and plausible reasoning* (Vol. 2). Princeton, NJ: Princeton University Press.

Researchware. (2011). HyperResearch v. 3.5.2. http://www.researchware.com. Accessed: February 2, 2016.

Schoenfeld, A. H. (Ed.). (1983). *Problem solving in the mathematics curriculum: A report, recommendations, and an annotated bibliography*. Washington, DC: Mathematical Association of America.

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

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

Skemp, R. R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching, 77,* 20–26.

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

Stylianides, G. J., & Silver, E. A. (2009). *Reasoning-and-proving in school mathematics: The case of pattern identification*. In D. A. Stylianou, M. L. Blanton, & E. J. Knuth (Eds.), *Teaching and learning proof across the grades: A K–16 perspective* (pp. 235–249). New York: Routledge.

Stylianou, D. A., Blanton, M. L., & Rotou, O. (2015). Undergraduate students' understanding of proof: Relationships between proof conceptions, beliefs, and classroom experiences with learning proof. *International Journal of Research in Undergraduate Mathematics Education, 1,* 91–134.

Thompson, D. R., & Senk, S. L. (2014). The same geometry textbook does not mean the same classroom enactment. *ZDM Mathematics Education, 46,* 781–795.

Chapter 15
About the Teaching and Learning of Proof and Proving: Cognitive Issues, Curricular Issues and Beyond

Lianghuo Fan and Keith Jones

Abstract In this commentary we provide an analytical look at the four studies reported within this theme of the volume and we discuss related issues and insights that we obtained from these studies. We offer a focused view of each study in the same order as presented in the four preceding chapters, and we conclude that these studies not only provide new insights into various cognitive and curricular issues in the teaching and learning of proof and proving, but also raise additional issues in relation to curriculum, textbooks and teacher education and professional development. Based on our analysis, we contend that mathematics education research in the area of proof and proving is still at an early stage, given that most studies are relatively small-scale and exploratory in nature. Further theoretical and methodological work, and more in-depth studies, especially larger-scale confirmatory and experimental studies, are needed to move research in this area forward.

Keywords Proof and proving · Cognition · Curriculum · Textbook analysis
Teacher education

Introduction

Issues relating to the teaching and learning of proof and proving (including related topics such as reasoning, explanation, argumentation and justification) have received mounting attention in research on mathematics education over the last two decades, and researchers have approached the issues from different perspectives in relation to pedagogy, curriculum, cognition, assessment and so on (see, e.g., Fan et al. 2017; Komatsu et al. 2017; Mariotti 2006; Miyazaki et al. 2017; Stylianides et al. 2017). The third theme of this volume consists of four chapters that address proof and proving with a particular focus on cognitive and curricular issues. In this commentary, we take each chapter in the sequence within the theme and discuss

L. Fan (✉) · K. Jones
University of Southampton, Southampton, UK
e-mail: l.fan@southampton.ac.uk

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_15

issues and insights we developed while engaging with these chapters. It should be noted that this commentary is not intended to present a comprehensive review of the chapters, nor to discuss in depth their strengths or weaknesses; rather the commentary positions the chapters within the research field and uses this as a springboard for considering what forms of future research would be most useful. The commentary closes with a summary of our more general observations and, based on our earlier considerations, some conclusions in relation to the theme.

Students' Learning of Proof and Proving

The first two chapters of this theme both focus on students' learning of proof and proving. In Askevold and Lekaus's chapter, Mathematical argumentation in Pupils' Written Dialogues, the researchers invited 33 fifth and sixth graders, aged 10–12, from two Norwegian primary classrooms, to write their own dialogues after they were given an introductory dialogue between two imaginary pupils about what is more desirable to be given, 1/10 or 1/3 of a cake. In this way, the participating pupils were expected to explain how eight given fractions 3/4, 2/3, 9/12, 3/7, 4/5, 4/8, 9/7, and 3/5 should be put in order from the least to the greatest. All pupils except one worked in groups of 2–4 to produce the dialogues. The researchers collected six dialogues from the fifth graders and 10 from the sixth graders. The results showed that while the sixth graders, compared with the fifth graders, used more rule-bound approaches based on conversions of fractions (60% vs. 50%), the fifth graders had a much higher percentage in using visual representation (100% vs. 10%) and in using both diagrammatic and narrative argumentation (50% vs. 30%).

We think it is meaningful to know how students at different grade levels perform in the area of proof and proving, in this case making mathematical argumentation, and, more importantly, why they perform differently and what might be the contributing factors to such difference. In Askevold and Lekaus's study, although the authors admit that the reasons for the difference were unclear, they argued that it might reflect the fact that students seem to note an expectation to use more formal mathematical language as they progress through school and that it might also reflect diverse teaching approaches applied in the two classrooms.

In our view, the difference revealed in Askevold and Lekaus's study fits with other research (e.g., Healy and Hoyles 2000) that, as students move to higher grade levels, they tend to think that their teachers expect abstract ways of representing mathematical concepts and that using more concrete and visual (pictorial/diagrammatic) representations is something that is used at a lower grade level (see also Leong et al. 2015). This is at odds with the professional practice of mathematicians where visual (pictorial/diagrammatic) representations are highly valued and frequently used (Burton 2004; Giaquinto 2007). This is an issue that warrants further investigation.

No doubt other influences play a part, as well. For example, the curriculum (including the textbooks) and, in particular, students' prior learning experiences, might well be relevant factors for explaining the differences found by Askevold and Lekaus. In this regard, it would be particularly interesting to know what learning experience (including home experience) might have helped the sixth grader who, uniquely, as reported in the study, used the method of finding a common numerator to compare the fractions. To tackle the issues of curriculum (including textbooks) and students' prior learning experiences may require different methods; for example, interview, curriculum and textbook analysis, and longitudinal study—all of these are undoubtedly more challenging.

Using interviews with a think-aloud protocol, Karunakaran's chapter, The Need for 'Linearity' of Deductive Logic: An Examination of Expert and Novice Proving Processes, compared the performances of five undergraduate mathematics students (whom he called "novice provers") with five advanced doctoral students majoring in mathematics (whom he called "expert provers") in their proving process with five novel mathematics statements in real analysis. He found that the advanced doctoral students were willing to knowingly, and temporarily, interrupt the deductive logic in their proving process, while the undergraduate students seemed less inclined to behave in a similar manner. The researcher used the empirical evidence obtained to argue that Expert and Novice provers approach the sequencing of 'bundles', which, according to the researcher, consist of groups of actions and resources that are clustered together by identifiable intentions (Karunakaran, this volume) in the proving process, in different ways. More specifically, expert provers showed more non-linearity while novice provers demonstrated more linearity in their proving process.

The issue which arose for us after reading Karunakaran's study was why there were these differences; the researcher did not elaborate on this issue, which may be explored in future research. The study divided mathematical provers into the novice and expert groups based on their stages of learning mathematics. However, many researchers, especially at the school level, have reported that students in the same learning stages or grade levels show remarkable differences in mathematical problem solving (e.g., Krutetskii 1976) and in geometric proving (e.g., Senk 1989; Usiskin 1982). In this regard, research could usefully focus on what are the possible contributing factors, and to what extent they contribute to such differences.

Karunakaran identified two implications of the results from the perspectives of curriculum developers, and undergraduate mathematics instructors; that many current undergraduate mathematics textbooks tend to present proof and proving in a linear style, and that undergraduate mathematics instructors may present proving as a sequential series of steps. Given that textbooks, especially traditional textbooks, are static and space or length-limited, and, moreover, that the presentation in textbooks needs to be concise and coherent, it remains a challenge how a textbook can be written or developed to present proofs to "reflect the nonlinear manner in which the proving process can occur and convey to students the idea that it is acceptable to assume the truth of a statement during the proving process as long as one eventually returns to address that assumption" (Karunakaran, this volume, p. 181). In this regard, a dynamic and interactive e-textbook could possibly offer new and better solutions about this issue (for papers on e-textbook development, as well as other related matters, see Jones et al. 2014). As to teachers/instructors, it seems to us that they can indeed play a larger part (compared with the textbooks) in providing students with experience of learning mathematics proofs in a suitably nonlinear manner. There is work to be done here in terms of research on teacher knowledge and professional development (Fan 2014).

Opportunities for Students to Learn 'Reasoning-and-Proving'

In the chapter, Reasoning-and-Proving in Algebra in School mathematics Textbooks in Hong Kong, Wong and Sutherland focused on the opportunities for students to learn 'reasoning-and-proving' from solving Algebra problems in a popular school mathematics textbook from Hong Kong. Here, the four major proof-related activities of 'reasoning-and-proving', as proposed by Stylianides (2009), comprising identifying mathematical patterns, making conjectures, providing non-proof arguments, and providing proofs, were adapted. Identifying mathematical patterns was sub-divided into plausible patterns and definite patterns; providing non-proof arguments was sub-divided into 'empirical argument' and 'rationale' (basically a statement of a result), while providing proofs was sub-divided into 'generic example' and Wong and Sutherland's adapted definition of 'demonstration' that included proof by definition and proof by calculation. For their analysis, Wong and Sutherland selected the algebra chapters (i.e., chapters in the 'Algebra and Number' strand) from the Hong Kong mathematics textbook for Year 10 students. Tasks involving reasoning-and-proving opportunities were coded as Type-1 (ones that explicitly asked for justification or explanation; their usual forms were "Prove that" and "Explain your answer"), Type-2 (ones that implicitly asked for justification; their usual form was "Determine whether …"), or Type-3 (what Wong and Sutherland called templates for illustrating reasoning-and-proving).

Wong and Sutherland found that of 3241 tasks in the algebra chapters, some 410 (i.e., 13%) designed opportunities for students to learn 'reasoning-and-proving'. What is more, they found almost no opportunities for conjecturing in the Algebra chapters and that identifying a pattern and providing support for proving were treated rather in isolation. Indeed, the majority of the reasoning-and-proving opportunities (280 out of 364, i.e., 77%) were classified as 'demonstration' (i.e., proof by definition and proof by calculation). Wong and Sutherland concluded that their analysis confirmed, to some extent, the findings of previous studies of secondary school classrooms in Hong Kong that student activities mostly focus on practicing and memorizing mathematical concepts and procedures (e.g., Leung 2001) and that their findings fit with the insight that, in international comparisons, Hong Kong students generally do better in 'knowing' than in 'reasoning' (e.g., Leung 2015). Needless to say, as Fan et al. (2013) have argued, there are many factors influencing students' learning. In studies about textbooks, the issues of whether the selected textbooks are a good representation of all the available textbooks for students to use, and whether the students whose academic performances were compared did use the textbooks, cannot be underplayed or taken for granted. We note further that, while all TIMSS studies (except TIMSS 1995) involved only fourth and eighth grade students, the textbook analyzed in the study of Wong and Sutherland is for tenth grade students. Clearly there is a gap in the research evidence and, we think, more research with a different research methodology is needed to draw a more confirmative conclusion about whether there exists any direct causal relationship between the textbooks and the students' achievements.

Wong and Sutherland did, of course, point out that 'reasoning-and-proving' does not, in general, always feature prominently in the Algebra component of the school mathematics curriculum. This can be despite the development of mathematical reasoning being one of the overall aims of the mathematics curriculum. Here the findings of Wong and Sutherland are not only in line with the findings of previous studies of secondary mathematics in Hong Kong but fairly in line with other analyses of 'reasoning-and-proving' opportunities in Algebra textbooks (or textbook sections) from other countries. For example, in an analysis of two US algebra textbooks, Davis et al. (2014) found that very few definite pattern opportunities were tied to the development of mathematical arguments and that conjecturing did not appear as frequently as other 'reasoning-and-proving' activities. This contrasts with analyses of school geometry textbooks (or textbook sections) by, for example, Fujita and Jones (2014) in the case of Japan, and Otten et al. (2014a, b) in the case of the USA, which generally found a much higher proportion of 'reasoning-and-proving' activities in the respective mathematics textbooks (or textbook sections). This raises the question of how, in the design of the school mathematics curriculum, the opportunities for students to learn 'reasoninhave revealed thatg-and-proving' are best developed across the curricular topics of number, algebra, geometry and stochastics. Moreover, as Fan (2013) pointed out, textbook analysis and comparison can primarily tell us how a particular topic is treated in the same series of textbooks, or how it is treated differently across different series of textbooks, but, without further evidence, it cannot go beyond this to tell us how the treatment of the topic should be improved and, in the case of textbook comparison, which textbook is better. Taking the Wong and Sutherland study as an example, if 13% of the 3241 tasks on 'reasoning-and-proving' is not adequate, then a further and challenging question is what percentage would be adequate and how this proportion might be justified.

In the final chapter of this theme, Irish Teachers' Perceptions of Reasoning-and-Proving amidst a National Educational Reform, Davis reported the perceptions of Irish mathematics teachers about pattern identification, conjecturing, and proof from mathematical, psychological and pedagogical perspectives based on Stylianides' conceptual framework of 'reasoning-and-proving' (Stylianides 2008). The data were collected from 10 male and 12 female teachers with different teaching experiences through semi-structured interviews lasting from 30 to 60 min. The results showed that the participating teachers perceived pattern identification and conjecturing as disconnected from proof construction, and, moreover, only seven teachers considered proof as connected to understanding. According to the researcher, while ‘reasoning-and-proving’ processes appear in the school mathematics syllabus in Ireland, teachers were not necessarily cognizant of this component of the syllabus as it is not reflected in specific mathematics content strands (in other words, the process and content strands were not interwoven).

The need for greater teacher awareness, and for greater professional development, is evident from the results of many studies of teachers’ knowledge in mathematics that have revealed that “teachers’ knowledge is very insufficient in quantity and unsatisfactory in quality” (Fan 2014, p. 36). As such, the findings reported in Davis’ study, though with a particular focus on ‘reasoning-and-proving’ in the Irish education context, are, in general, consistent with those of many other studies.

As well as agreeing with Davis that part of the problem about teachers’ inadequate knowledge of ‘reasoning-and-proving’ can be alleviated by improving the curriculum design, we think his study highlights the important issue of teacher education and professional development. Specifically, Davis’ study raises the issue of how pre-service teacher education, in-service teacher education, and other teacher professional development activities can help prospective and current teachers to develop adequate knowledge of mathematics, and, in this particular case, of reasoning and proof. Further research in this direction would be highly valuable.

Concluding Remarks

We are pleased to see that the teaching and learning of proof and proving, being an important topic, continues to receive increasing attention from mathematics education researchers internationally, as evidenced in this theme as well as this volume. We are encouraged to notice that different researchers have collectively shown the breadth and depth of the research in this area, and not only have they provided new insights into various aspects of the teaching and learning of proof and proving, with a focus on cognitive and curricular issues in particular, but, through their work, also raised new issues in relation to curriculum, textbooks and teacher education and professional development.

Nevertheless, it can be seen that mathematics education research in this area is, overall, still at an early stage, given that most, if not all, of the studies are of small-scale and exploratory in nature. It is clear to us that more studies, in particular ones that are large-scale, confirmatory and experimental in nature, are needed to develop this area of research further. We are sure that the studies reported in this theme, together with other studies in this volume, can serve as an important foundation for the international research community to move forward in this valued area of research in mathematics education.

References

Burton, L. (2004). *Mathematicians as enquirers: Learning about learning mathematics*. Dordrecht, The Netherlands: Kluwer.

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

Fan, L. (2013). Textbook research as scientific research: Towards a common ground on issues and methods of research on mathematics textbooks. *ZDM-International Journal on Mathematics Education, 45*(5), 765–777.

Fan, L. (2014). *Investigating the pedagogy of mathematics: How do teachers develop their knowledge?* London: Imperial College Press.

Fan, L., Qi, C., Liu, X., Wang, Y., & Lin, M. (2017). Does a transformation approach improve students' ability in constructing auxiliary lines for solving geometric problems? An intervention-based study with two Chinese classrooms. *Educational Studies in Mathematics*, *96*(2), 229–248.

Fan, L., Zhu, Y., & Miao, Z. (2013). Textbook research in mathematics education: Development status and directions. *ZDM-International Journal on Mathematics Education, 45*(5), 633–646.

Fujita, T., & Jones, K. (2014). Reasoning-and-proving in geometry in school mathematics textbooks in Japan. *International Journal of Educational Research, 64,* 81–91. https://doi.org/10.1016/j.ijer.2013.09.014.

Giaquinto, M. (2007). *Visual thinking in mathematics*. Oxford: Oxford University Press.

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

Jones, K., Bokhove, C., Howson, G., & Fan, L. (Eds.). (2014). *Proceedings of the International Conference on Mathematics Textbook Research and Development* (*ICMT-2014*). Southampton: University of Southampton. https://eprints.soton.ac.uk/374809/.

Komatsu, K., Jones, K., Ikeda, T., & Narazaki, A. (2017). Proof validation and modification in secondary school geometry. *Journal of Mathematical Behavior*, *47*, 1–15. https://doi.org/10.1016/j.jmathb.2017.05.002.

Krutetskii, V. A. (1976). *The psychology of mathematical abilities in schoolchildren*. Chicago: University of Chicago Press.

Leong, Y. H., Ho, W. K., & Cheng, L. P. (2015). Concrete-pictorial-abstract: Surveying its origins and charting its future. *The Mathematics Educator, 16*(1), 1–19.

Leung, F. K. S. (2001). In search of an East Asian identity in mathematics education. *Educational Studies in Mathematics, 47*(1), 35–51.

Leung, F. K. S. (2015). What does TIMSS inform us about mathematics (under)achievers in Hong Kong? In K. L. Wong, et al. (Eds.), *Proceedings of the Hong Kong Mathematics Education Conference 2015*. Hong Kong: Hong Kong Association for Mathematics Education.

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

Miyazaki, M., Fujita, T., & Jones, K. (2017). Students' understanding of the structure of deductive proof. *Educational Studies in Mathematics, 94*(2), 223–239. https://doi.org/10.1007/s10649-016-9720-9.

Otten, S., Gilbertson, N. J., Males, L. M., & Clark, D. L. (2014a). The mathematical nature of reasoning-and-proving opportunities in geometry textbooks. *Mathematical Thinking and Learning, 16*(1), 51–79.

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

Senk, S. L. (1989). Van Hiele levels and achievement in writing geometry proofs. *Journal for Research in Mathematics Education, 20*(3), 309–321.

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

Stylianides, G. J. (2009). Reasoning-and-proving in school mathematics textbooks. *Mathematical Thinking and Learning, 11*(4), 258–288.

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

Usiskin, Z. (1982). *Van Hiele levels and achievement in secondary school geometry*. Chicago, IL: CDASSG Project, University of Chicago. Available online at http://ucsmp.uchicago.edu/resources/van-hiele/. Accessed July 5, 2017.

Part IV
Issues Related to the Use of Examples in Proof and Proving

Chapter 16
How Do Pre-service Teachers Rate the Conviction, Verification and Explanatory Power of Different Kinds of Proofs?

Leander Kempen

Abstract In the opening session of a course for first-year secondary (lower track secondary school) pre-service teachers, the participants were asked to rate the conviction, verification and explanatory power of four different kinds of proofs (a generic proof with numbers, a generic proof in the context of figurate numbers, a proof in the context of figurate numbers using "geometric variables" and the formal proof). In this study, students' ratings express their preference for the formal proof concerning the aspects conviction, verification, and explanatory power. The other proofs achieve significantly lower ratings, especially in the case of conviction. The results may open the discussion about the use of generic proofs, the use of figurate numbers and the concept of proofs that explain.

Keywords Transition to university · Generic proof · Figurate numbers
Function of proof

Introduction

The University of Paderborn requires the course "Introduction into the culture of mathematics" for all first-year pre-service teachers (lower track secondary school) to help them to accomplish the transition to higher mathematics. This course has been developed and taught by Biehler and Kempen (2013). Refining and evaluating the course is a main focus of the author's dissertation. In this course, four different kinds of proofs are used to foster students' proof skills: the generic proof with numbers, the generic proof in the context of figurate numbers, the formal proof and the proof in the context of figurate numbers using "geometric variables" (Kempen and Biehler 2016). The course's three main objectives are: (1) to enhance students' transition to the mathematical formal proof, (2) to promote the mathematical

L. Kempen (✉)
University of Paderborn, Paderborn, Germany
e-mail: kempen@khdm.de

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_16

symbolic language in a meaningful way and (3) to equip students with "intellectual-honest" (Stylianides 2007) ways of proving, that can be used in school mathematics later on. The course is evaluated and refined in a design based research scenario (Gravemeijer and Cobb 2006). In this chapter I will outline the part of the project in which students rated the conviction, verification and explanatory power of the different kinds of proofs in the beginning of the course.

Theoretical Background

In the teaching of mathematical proof, different kinds of proofs have been introduced and discussed by mathematics educators (Dreyfus et al. 2012). In various domains of education much attention has been paid to the concept of generic proof (e.g., Rowland 2002; Karunakaran et al. 2014; Stylianides 2010); "A generic proof aims to exhibit a complete chain of reasoning from assumptions to conclusion, just as in a general proof; however, […] a generic proof makes the chain of reasoning accessible to students by reducing its level of abstraction; it achieves this by examining an example that makes it possible to exhibit the complete chain of reasoning without the need to use a symbolism that the student might find incomprehensible" (Dreyfus et al. 2012, p. 204). From a pedagogical point of view, several important questions arise when using a generic proof: How can one expose the generality of an argumentation given in a concrete context to a reader of the proof? How can a reader of a 'generic proof' know what part of the concrete examples presented are meant to be generic? (cf. Biehler and Kempen 2013; Mason and Pimm 1984; Reid and Vallejo Vargas 2017, this volume). Building on this discussion, Biehler and Kempen (2013) developed a pedagogical concept of generic proofs: In a generic proof the generic argument is illustrated in concrete examples and its validity and generality is explicitly expressed in words. In the following, I will refer to this concept. (Examples of generic proofs will be given below.)

In contrast to the suggestions in the literature for the use of the generic proofs, their benefits and usefulness for the learning of mathematical proof has not been investigated in detail yet. With this research study, I want to contribute to the ongoing discussion on the usefulness of different kinds of proofs with respect to conviction, verification and explanation.

In mathematics proofs are said to cover different functions. The most prominent ones are verification/conviction, explanation, systematization, discovery and communication (e.g., de Villiers 1990). Referring to the work of Hersh (1993) I argue that there are two important aspects of proofs: conviction and explanatory power. In dealing with the function conviction, the distinction between relative and absolute conviction made by Weber and Mejia-Ramos (2015) seems to be fruitful to clarify two different functions a proof might fulfill. Absolute conviction is about the mathematical 'objective' truth of a statement, verified by a mathematical proof. In the following, this function of establishing objective truth will be called *verification*. Relative conviction is meant as a personal subjective conviction that a statement is or might be true: a reader of a proof "has a relative conviction in a claim if the subjective level of probability that one attribute to that claim being true exceeds a certain" (ibid., p. 16). This relative conviction will be meant when using the term *conviction* in the following. An important distinction in the teaching of proofs has been given by Hanna (1989) who distinguishes between proofs that (only) prove that a theorem is true and proofs that (also) explain why it is true. From an educational point of view she highlights the function of *explanation*. In the following part, I will give a brief theoretical outline on these, as I consider three different functions of proofs: verification, conviction and explanation.

Verification

Verification is concerned with the 'truth' of a statement. A proof verifies a statement by showing that it is a necessary conclusion from axioms or previous proved theorems. Everyone who agrees with the arguments and logic inferences has to agree with the concluding results. It is this view on mathematical proof that lends the finding to be timeless (Weber 2014): The proven theorem will be 'true' forever.

Conviction

Mathematical proof is said to be a convincing argument (e.g., Hanna 1989; Hersh 1993). Here, conviction is considered a personal and subjective category (see above). A proof may convince us that a statement is true, i.e., the reader is persuaded without any doubt that the statement holds in every possible case and that no counterexample may exist. This view on conviction is related to the concept of "epistemic value" of Duval (1990, 2007). As Reid and Knipping (2010, p. 74) mention: "it is important to recognize that while logically a statement can only be true or false, psychologically it can take on one of many values, which Duval (1990, 2007) calls its "epistemic value"". Accordingly, the epistemic value highlights the individual perception of conviction as a personal judgement of whether a proposition is believed (cf. ibid., p. 74).

Explanation

Hanna (1989) stresses that a proof can give insight as to why a statement is true. She further states: "I will say that proof explains when it shows what "characteristic property" entails the theorem it purports to prove" (Hanna 1989, p. 47). Explanatory proofs often make use of geometric descriptions to reach the conclusion. These kinds of representations are said to be more comprehensible or accessible for learners and ease the transition to algebra (compare Flores 2002). As explicated by Hanna (2017, in this book), explanation can have different meanings. In the philosophy of mathematics an intra-mathematical focus on explanation is emphasized, stressing the connections between mathematical statements and their mutual relationships (compare the quotation by Hanna above). From a pedagogical point of view, explanation can be understood as conveying some kind of insight, why a mathematical statement is true. In this sense, explanation is closely related to the aim of personal understanding. However, for this study, one has to stress that ‘to explain’ is defined implicitly by what is meant individually by the students.

Following the theoretical considerations above, verification can be considered a concept that stands for its own. A reader of a proof might perceive the necessity of a conclusion in a proof or not. However, one might identify a link between explanation and (relative) conviction. Conviction is linked to a personal judgement based on one’s knowledge and understanding. The initial point of an explanation is a deficit in someone’s knowledge. Accordingly, an explanation aims at increasing someone’s knowledge and thus may change his epistemic situation (compare Kiel 1999, p. 72 f.) and therefore may lead to conviction.

Findings from the Literature

Results from different studies suggest that some students do not accept deductive proofs as verification (e.g., Fischbein and Kedem 1982; Healy and Hoyles 2000; for an overview, see Reid and Knipping 2010). In these studies, after having seen a correct deductive proof, some learners did not accept the immanent general verification of the proof to cover all possible cases. This non-acceptance of a correct proof gets even more important when looking at generic proofs. In the study of Tabach et al. (2010) about half of the secondary school teachers rejected correct generic proofs due to a perceived lack of generality. In the study of Martin and Harel (1989) between 42 and 46% of the 101 pre-service teachers gave only low ratings to the generic proofs (‘particular proof’) concerning verification. Also Dreyfus (2000) and Knuth (2002) showed that teachers might underestimate proofs making use of concrete examples and narratives. Kempen and Biehler (2016) identified different perceptions of generic proofs (with numbers) and found that few students were convinced by the generic proofs both from a logical and psychological perspective.

In the study of Healy and Hoyles (2000), students were asked to rate different types of arguments (empirical, algebraic and narrative) with regard to explanatory power (the answer “…is an easy way to explain to someone in your class who is unsure”; ibid., p. 403). There, the algebraic arguments had the lowest ratings concerning explanatory power, whereas the narrative arguments obtained the highest. Concerning students’ preference for their own approach to prove a theorem, the authors conclude: “students preferred arguments that they could evaluate and that they found convincing and explanatory, preferences that excluded algebra” (ibid., p. 426).

Research Questions

This chapter focuses on the degree of conviction, verification and explanatory power perceived by pre-service teachers when reading different kinds of proofs. The research question is: How do pre-service teachers (lower track secondary school) rate the conviction, verification and explanatory power of four different kinds or proofs (the generic proof with numbers, the generic proof in the context of figurate numbers, the proof in the context of figurate numbers using "geometric variables", and the formal proof) at the beginning of a course for first-year students?

This research question is a part of a wider research project, where the impact of the course "Introduction into the culture of mathematics" was evaluated with a pre- and a post-test. The focus of this chapter is on the results of the pre-test that took place in the first session of the course. The results also give insight into students' understanding of mathematical proof when entering university.

Methodology

In the first session of the course, the participants were asked to complete a proof questionnaire (paper and pencil). These students' had passed the German Abitur (final secondary school examination) when graduating at the 'Gymnasium' (higher track secondary school). The questionnaire included each type of concrete proof mentioned above, one type for each statement (see below). The students had to rate different aspects of the proofs on a six-level Likert scale ([1] totally disagree … [6] totally agree). The aspects to be rated were verification, generality, conviction, explanatory power, and acceptance as correct and valid proof. It was then possible to construct a high reliable scale of "proof acceptance". In this contribution, the focus is on the questionnaire's following three statements: (i) "The reasoning convinces me that the statement is true" [conviction]; (ii) "The reasoning shows that the statement is true for every time and 100%" [verification]; (iii) "The reasoning explains to me why the statement is true" [explanatory power]. In addition to the following proofs, no further information was given to the students. The items to be answered (see above) did not contain the annotations "conviction", "verification" and "explanation".

I chose a different statement for each proof being rated. When using four different proofs to one statement in a previous pilot study, I identified influences between the different proofs that weakened the results. So I made the choice for using four different statements, even though in this case the ratings of one proof might be influenced by the correspondent statement. The proofs to be rated in this study were selected after piloting different kinds of proofs to different statements twice.

The students answered the questionnaire by using an anonymous code. Their answers had no impact on students' grades in the course in any way. The proofs to be rated are shown below.

Statement (1): The sum of an odd natural number and its double is always odd.

Generic proof with numbers:

$$1 + 2 \cdot 1 = 3 \cdot 1 = 3, \quad 5 + 2 \cdot 5 = 3 \cdot 5 = 15, \quad 13 + 2 \cdot 13 = 3 \cdot 13 = 39$$

The sum of an odd natural number and its double equals three times the initial number. Since the initial number is an odd number, one obtains the product of two odd numbers. Since the product of any two odd numbers is always odd, the result will always be an odd number.

Statement (2): The sum of five consecutive natural numbers is always divisible by five.

Generic proof in the context of figurate numbers:
In the representation of the sum of five consecutive natural numbers by figurate numbers, one always obtains the same shape of stairs on the right side. By transforming these stairs—taking the edge at the bottom right and putting it above—one always obtains five equal rows. So the result will always be divisible by five (Fig. 16.1).

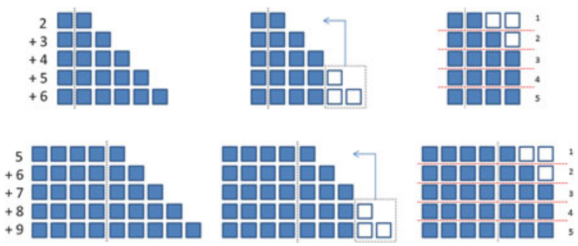


Fig. 16.1 The sum of five consecutive numbers represented by figurate numbers

Statement (3): The square of an even natural number is always divisible by four.

Proof with geometric variables:
(This proof was given without any verbal explanation. Here, the geometric variables are used to express the generality that has to be explained in the case of the generic proofs above.) (Fig. 16.2).

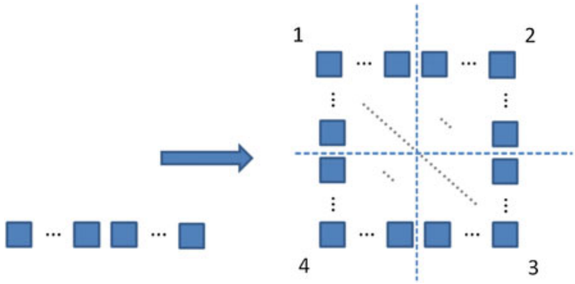


Fig. 16.2 A proof with "geometric variables" and figurate numbers

Statement (4): For all natural numbers a, b, c: If b is a multiple of a and c is a multiple of a, then $(b + c)$ is a multiple of a.

Formal Proof:

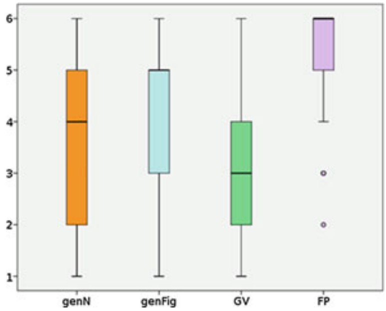
Let a, b, c be natural numbers. Since b is a multiple of a, there exists a natural number n with: $n \cdot a = b$. Since c is a multiple of a, there exists a natural number m with: $m \cdot a = c$. We get: $b + c = n \cdot a + m \cdot a = (n + m) \cdot a$. Since $(n + m)$ is a natural number, $(b + c)$ is a multiple of a. □

Results

In our study 149 pre-service teachers (94 female and 55 male; age: $\bar{a} = 21.14$) were asked to rate four different kinds of proofs: the generic proof with numbers ("GenN"), the generic proof in the context of figurate numbers ("GenFig"), the formal proof ("FP"), and the proof in the context of figurate numbers using "geometric variables" ("GV"). Here, I am examining on the proofs' conviction, verification and explanatory power. Each aspect was rated on a six-level Likert scale ([1] totally disagree … [6] totally agree). The results of students' ratings are shown below. (Since not all students have answered every question, some results refer to a sample size less than 149.)

Conviction (ratings of the item "The reasoning convinces me that the statement is true.")

Students rated the proof using geometric variables as the proof that was the least convincing with a median of 3. The generic proof with numbers had a median of 4, which means "just a little agreement", but the responses show a higher variation (first quartile: 2, third quartile 5). The generic proof with figurate numbers (median: 5) and the formal proof (median: 6) were rated the highest (see Fig. 16.3). The



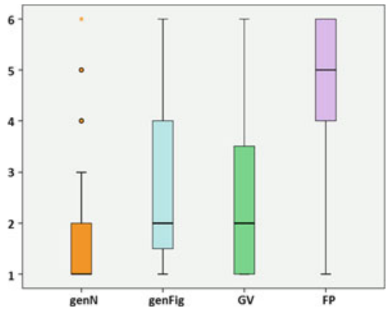
conviction	GenN	GenFig	GV	FP
n	145	146	138	145
mean	3.41	4.25	3.02	5.48
median	4.00	5.00	3.00	6.00
SD	1.66	1.50	1.67	.92

Fig. 16.3 Boxplots and statistical data concerning the item "conviction" ([1] totally disagree … [6] totally agree)

differences concerning the medians are pairwise highly statistically significant ($p < 0.001$), the difference between the medians of "GenN" (4) and "GV" (3) is significant with 'only' $p = 0.036$ (Wilcoxon-test).

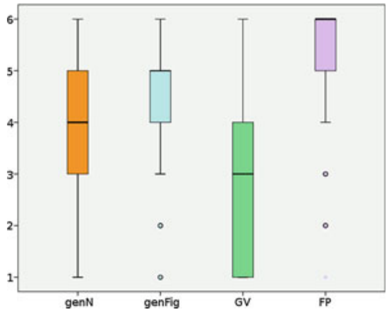
Verification (ratings of the item "The reasoning shows that the statement is true for every time and 100%.")

Concerning verification, the generic proof with numbers had the median of 1, which means "totally disagree". Having a look at the boxplot (see Fig. 16.4), the position of the box seems to be considerable (first quartile: 1, third quartile: 2). The generic proof with figurate numbers and the proof with geometric variables had a median of 2, but the responses show a high variation. The formal proof (median: 6) was rated the highest (first quartile: 4, third quartile: 6). All differences concerning the medians are pairwise highly statistically significant ($p < 0.001$; Wilcoxon-test).



verification	GenN	GenFig	GV	FP
n	147	146	137	145
mean	2.01	2.79	2.48	4.88
median	1.00	2.00	2.00	5.00
SD	1.40	1.63	1.56	1.37

Fig. 16.4 Boxplots and statistical data concerning the item "verification" ([1] totally disagree … [6] totally agree)



explan.	GenN	GenFig	GV	FP
n	147	145	138	145
mean	3.74	4.39	2.96	5.34
median	4.00	5.00	3.00	6.00
SD	1.51	1.36	1.74	1.09

Fig. 16.5 Boxplots and statistical data concerning the item "explanatory power" ([1] totally disagree [6] totally agree)

Explanatory power (ratings of the item "The reasoning explains why the statement is true.")

With regard to explanatory power, the proof with geometric variables had the lowest median. In this case, the high variation of ratings is remarkable (first quartile: 1, third quartile: 4). The generic proof with numbers was rated with a median of four and the generic proof with figurate numbers with a median of five. Concerning the explanatory power, the formal proof is considered the best (see Fig. 16.5). All differences concerning the median are pairwise highly statistically significant ($p < 0.001$; Wilcoxon-test).

Discussion

In this study, 149 pre-service teachers were asked to rate four different kinds of proofs (a generic proof with numbers, a generic proof in the context of figurate numbers, a proof in the context of figurate numbers using "geometric variables", and the formal proof) concerning the aspects conviction, verification and explanation. The study took place in the first session of a mathematics course for first-year students.

At the beginning of the course, the formal proof achieved the highest ratings in all the three categories: "conviction", "verification" and "explanatory power". Concerning conviction, the generic proofs got the medians of 4 and 5, whereas the proof with geometric variables was rated the lowest with a median of 3. With regard to verification, the generic proofs were rejected by most of the students (median of 1 and 2), as was the proof with geometric variables (median of 2). These results also show that conviction and verification can be distinguished in this study. Most of the students stated little agreement to the explanatory power of the generic proofs (medians of 4 and 5). And the proof with geometric variables was only considered slightly explanatory (median of 3) with high variation.

Comparing the generic proofs, the generic proof in the context of figurate numbers is rated higher in all three cases (conviction, verification and explanation). Accordingly, the notational system of figurate numbers seems to be more convincing and explanatory for the students and seems to imply a higher form of verification for them.

The fact that the formal proof is always rated higher than the generic proofs is a little surprising, because in the generic proofs the argument is explicitly written down, in addition to the concrete examples that illustrate the argument. These results conflict with the assumption that generic proofs are more explanatory by themselves (compare Hemmi 2006; Rowland 1998). But one can conclude that in this study, the mathematical symbolic language is perceived as both convincing and explanatory by the students and that the pre-service teachers accept the verification fulfilled in the formal proof.

Concerning the proof with geometric variables, one might assume that this kind of representation might not be known to all students. In addition, this proof was the only one that was not accompanied by any narrative or algebraic-symbolic expression. This fact might also explain the high variation of ratings.

It seems obvious that students' ratings of the proofs also rely on their former mathematics classes at school. Here, the concept of the didactical contract of Brousseau (1997) and the theory of socio-mathematical norms of Yackel and Cobb (1996) can be taken into account to give explanations for students' choices. Students learned explicitly and implicitly what kind of argument they might consider as a proof or not during their former mathematics classes at school. In addition, the content of an argument and its appearance has to be considered. Also Healy and Hoyles (2000) and Stylianides and Stylianides (2009) stress that the appearance of an argument influences learners' evaluation.

Following these considerations, the results do not only depend on the different kinds of proofs (the appearance) and on the different statements being proved, but also on the different educational background of the participants. However, as has been shown above, a geometrical representation does not guarantee that the proof will be considered explanatory. On the contrary, when the students are not familiar with this kind of representation, they may not understand the argument. These findings stress the awareness that a representation is neither self-evident nor self-explanatory and that its use does not necessarily lead to understanding (cf. Jahnke 1984). These results can be supported by taking a semiotic perspective. I.e., Peirce introduces the term 'collateral knowledge' to subsume all the knowledge one needs to read and to work in a notational system with some kind of representations he calls diagrams (Hoffmann 2005; Stjernfelt 2000). This collateral knowledge has to be developed by learners to do mathematics in general and to perform and to understand mathematical reasoning in particular. Following these considerations, learners must spent some time to acquire representations and symbols as mathematical tools. Working in and reading a notational system has to be learned explicitly (as proposed in the context of diagram literacy in Diezmann and English 2001). But as was shown above, when the mathematical symbolic language has been acquired by a learner, it becomes a convincing and explanatory tool to fulfill verification. More research is needed to investigate the role of students' former mathematics classes, the content of an argument and its representation when dealing with mathematical proof.

Having in mind that generic proofs and figurate numbers are said to be useful and adequate tools to perform reasoning and proving even in school mathematics, several questions arise: How can school students develop an adequate understanding of mathematical proof when they might not understand or accept the general verification fulfilled by a given (generic) proof either? Do school students or students at university have enough time to acquire collateral knowledge about all the representations and symbols the teachers want them to use? How might students at school and at university learn about the different functions of mathematical proof when not perceiving them? Following these considerations, it seems valuable to highlight the meaning of proof acceptance when discussing forms of proofs and proving in the classroom.

As Hanna (1995) points out, explanatory proof can have different forms depending on the classroom context and the experience of the learners. According to the results of this study, one has to consider the explanatory power of the mathematic symbolic language and to rethink the concept of proofs that prove and proofs that explain.

References

Biehler, R., & Kempen, L. (2013). Students' use of variables and examples in their transition from generic proof to formal proof. In B. Ubuz, C. Haser, & M. A. Mariotti (Eds.), *Proceedings of the 8th Congress of the European Society for Research in Mathematics Education* (pp. 86–95). Ankara: Middle East Technical University.

Brousseau, G. (1997). *Theory of didactical situations in mathematics: Didactique des mathématiques 1970–1990* (N. Balacheff, M. Cooper, R. Sutherland, & V. Warfield, Eds. & Trans.). Dordrecht: Kluwer Academic Publishers.

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

Diezmann, C., & English, L. (2001). Promoting the use of diagrams as tools for thinking. In A. Cuoco & F. R. Curcio (Eds.), *The roles of representation in school mathematics* (pp. 77–89). Reston, VA: National Council of Teachers of Mathematics.

Dreyfus, T. (2000). Some views on proofs by teachers and mathematicians. In A. Gagatsis (Ed.), *Proceedings of the 2nd Mediterranean Conference on Mathematics Education* (Vol. 1, pp. 11–25). Nikosia: The University of Cyprus.

Dreyfus, T., Nardi, E., & Leikin, R. (2012). Forms of proof and proving in the classrom. In G. Hanna & M. de Villiers (Eds.), *Proof and proving in mathematics education: The 19th ICMI study* (pp. 191–214). Heidelberg: Springer Science+Business Media.

Duval, R. (1990). Pour une approche cognitive de I'argumentation. *Annales de Didactique et de Sciences Cognitives, 3,* 195–221.

Duval, R. (2007). Cognitive functioning and the understanding of mathematical processes of proof. In P. Boero (Ed.), *Theorems in school: From history, epistemology and cognition to classroom practice* (pp. 137–161). Rotterdam: Sense Publishers.

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

Flores, A. (2002). Geometric representations in the transition from arithmetic to algebra. In F. Hitt (Ed.), *Representation and mathematics visualization* (pp. 9–30). North American Chapter of the International Group for the Psychology of Mathematics Education.

Gravemeijer, K., & Cobb, P. (2006). Design research from the learning design perspective. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational Design research: The design, development and evaluation of programs, processes and products* (pp. 45–85). London: Routledge.

Hanna, G. (1989). Proofs that prove and proofs that explain. In G. Vergnaud, J. Rogalski, & M. Artigue (Eds.), *Proceedings of the 13th International Conference on the Psychology of Mathematics Education* (Vol. 2, pp. 45–51). Paris: PME.

Hanna, G. (1995). Challenges to the importance of proof. *For the Learning of Mathematics, 15*(3), 42–49.

Hanna, G. (2017). Reflections on proof as explanation. *Proceedings of the 13th International Congress on Mathematical Education*.

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education, 31*(4), 396–428.

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

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

Hoffmann, M. (2005). *Erkenntnisentwicklung*. Frankfurt am Main: Klostermann.

Jahnke, H. N. (1984). Anschauung und Begründung in der Schulmathematik *Beiträge zum Mathematikunterricht* (pp. 32–41). Bad Salzdethfurt.

Karunakaran, S., Freeburn, B., Konuk, N., & Arbaugh, F. (2014). Improving preservice secondary mathematics teachers' capability with generic example proofs. *Mathematics Teacher Educator, 2*(2), 158–170.

Kempen, L., & Biehler, R. (2016). Pre-service teachers' perceptions of generic proofs in elementary number theory. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the 9th Congress of the European Society for Research in Mathematics Education* (pp. 135–141). Prague: CERME.

Kiel, E. (1999). *Erklären als didaktisches Handeln*. Würzburg: Ergon-Verlag.

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

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

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

Reid, D. A., & Knipping, C. (2010). *Proof in mathematics education. Research, learning and teaching*. Rotterdam: Sense Publisher.

Reid, D. A., & Vallejo Vargas, E. A. (2017). When is a generic argument a proof? In *Proceedings of the 13th International Congress on Mathematical Education*.

Rowland, T. (1998). Conviction, explanation and generic examples. In A. Olivier & K. Newstead (Eds.), *Proceedings of the 22nd Conference of the International Group for Psychology of Mathematics Education* (Vol. 4, pp. 65–72). Stellenbosch, South Africa: University of Stellenbosch.

Rowland, T. (2002). Generic proofs in number theory. In S. R. Campbell & R. Zazkis (Eds.), *Learning and teaching number theory* (pp. 157–183). Westport, Connecticut: Ablex.

Stjernfelt, F. (2000). Diagrams as centerpiece of a Peircean epistemology. *Transactions of the Charles S. Peirce Society, 36*(3), 357–384.

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

Stylianides, G. J. (2010). Engaging secondary students in reasoning and proving. *Mathematics Teaching, 219,* 39–44.

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

Tabach, M., Barkai, R., Tsamir, P., Tirosh, D., Dreyfus, T., & Levenson, E. (2010). Verbal justification—Is it a proof? Secondary school teachers' perceptions. *International Journal of Science and Mathematics Education, 8*(6), 1071–1090.

Weber, K. (2014). Proof as a cluster concept. In C. Nicol, S. Oesterle, P. Liljedahl, & D. Allan (Eds.), *Proceedings of the 38th Conference of the International Group for the Psychology of Mathematics Education and the 36th Conference of the North American Chapter of the Psychology of Mathematics Education* (Vol. 5, pp. 353–360). Vancouver, Canada: PME.

Weber, K., & Mejia-Ramos, J. P. (2015). On relative and absolute conviction in mathematics. *For the Learning of Mathematics, 35*(2), 15–21.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*(4), 458–477.

Chapter 17
When Is a Generic Argument a Proof?

David Reid and Estela Vallejo Vargas

Abstract We discuss whether a generic argument can be considered a proof. Two positions on this question have recently been published which focus on the fussiness of an argument as a deciding criterion. We take a third view that takes into account psychological and social factors. Psychologically, for a generic argument to be a proof it must result in a convincing deductive reasoning process occurring in the mind of the reader. Socially, for a generic argument to be a proof it must conform to the social conventions of the context. For classroom settings, we suggest two kinds of evidence that should be reflected in written work in order for a generic argument to be accepted as a proof. These kinds of evidence reveal the linkage between the psychological and social factors.

Keywords Generic arguments · Proof · Social perspectives · Psychological perspectives · Evidence

Introduction

A generic argument "involves making explicit the reasons for the truth of an assertion by means of operations or transformations on an object that is not there in its own right, but as a characteristic representative of its class. The account involves the characteristic properties and structures of a class, while doing so in terms of the names and illustration of one of its representatives" (Balacheff 1988, p. 219). The "characteristic representative of its class" is called a *generic example* (Mason and Pimm 1984). Such arguments have been discussed in the mathematics education literature since Mason and Pimm introduced the terms "generic example" and

D. Reid (✉)
University of Bremen, Bremen, Germany
e-mail: dreid@math.uni-bremen.de

E. Vallejo Vargas
Pontifical Catholic University of Peru, Lima, Peru
e-mail: vallejo@uni-bremen.de

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_17

"generic proof", and this discussion has frequently included debate about if and when generic argument are acceptable mathematical proofs. For example, Leron and Zaslasvsky (2013) write,

> The main weakness of a generic proof is, obviously, that it does not really prove the theorem. The "fussiness" of the full, formal, deductive proof is necessary to ensure that the theorem's conclusion infallibly follows from its premises. (p. 27)

Yopp et al. (2015) contest this,

> We … propose that students can make and judge the viability of generic example arguments and that in certain situations … these arguments can be accepted as proof. (p. 10)

In this chapter we take up this debate, and argue that the main criterion used by Leron and Zaslavsky as well as Yopp, Ely, and Johnson-Leung to distinguish proofs from non-proofs, "fussiness", is not in fact the criterion that is critical in determining whether an argument using a generic example is a mathematical proof. We take an alternative view that takes into account also psychological and social factors.

Fussiness

Movshovitz-Hadar (1988) also maintains that a generic argument cannot be a mathematical proof.

> The proof of a generic example should not be confused with a fully general proof. It only suggests the full proof through a generalizable concrete example. From the purely logical point of view there is no replacement for the formal proof. (p. 18)

In mathematics education there is wide agreement that formal proofs are needed. The NCTM (2000) Standards say that students should understand that a proof is an argument "consisting of logically rigorous deductions of conclusions from hypotheses" (NCTM 2000, p. 56). The need for logically rigorous deductions based on previously established propositions is the "fussiness" called for by Leron and Zaslavsky and which Yopp, Ely, and Johnson-Leung claim generic argument can achieve "in certain situations".

But is fussiness actually a characteristic of mathematical proofs? Not if it is absolute and complete fussiness.

> For many mathematical investigations, full mathematical formalization and complete formal proof, even if possible in principle, may be impossible in practice. They may require time, patience, and interest beyond the capacity of any human mathematician. Indeed, they can exceed the capacity of any available or foreseeable computing system. (Hersh 1993, p. 390)

Aberdein (2012) elaborates Epstein's (2012) model of proof based on mathematical practice. He characterizes mathematical proofs as an argument with two parallel structures, one argumentational and the other inferential. The inferential structure has absolute and complete fussiness. Every step is a deduction based on and justified by previous propositions. It is fully formal. But it is never actually presented. As Hersh points out this is usually impossible. Instead, using the argumentational structure "mathematicians attempt to convince each other of the soundness of the inferential structure" (Aberdein 2012, p. 362).

> This account both conserves and transcends the conventional view of mathematical proof. The inferential structure is held to strict standards of formal rigour, without which the proof would not qualify as mathematical. However, the step-by-step compliance of the proof with these standards is itself a matter of argument, and susceptible to challenge. Hence much actual mathematical practice takes place in the argumentational structure. (p. 363)

mathematical proofs are not "logically rigorous deductions of conclusions from hypotheses" as the NCTM asserts. Instead, they are arguments that such deductions exist. As Hardy observed long ago, proofs do not prove in the formal sense, they point.

> If we were to push it to its extreme we should be led to a rather paradoxical conclusion; that we can, in the last analysis, do nothing but *point*; that proofs are what Littlewood and I call gas, rhetorical flourishes designed to affect psychology, pictures on the board in the lecture, devices to stimulate the imagination of pupils. (Hardy 1928, p. 18, his emphasis)

Furthermore proofs are not completely fussy.

> Mathematical arguments, just like arguments in our daily lives, leave much unsaid. And of what is said, much is only hints or sketches, with lots explicitly left to the reader. (Epstein 2012, p. 269)

Psychological Factors

Whether or not generic argument are fussy, we believe this is not the criterion that determines if they are proofs, as proofs are not completely fussy either. To be proofs, generic argument must fulfill the function of argumentative structures, to *point* to the inferential structure, to affect psychology, to stimulate the imagination. As Fischbein (1982) states, "there are frequent situations in mathematics in which a formal conviction, derived from a formally certain proof, is NOT associated with the subtle feeling of 'It must be so', 'I feel it must be so'" (p. 11). So, for some readers this stimulation of the reader's mind might be more difficult to achieve using fussy formal proofs than with generic arguments that qualify as proofs.

As examples of generic argument, we include the three following arguments for the claim: "Prove that the sum of the first n natural numbers is $\frac{n(n+1)}{2}$".

Argument 1

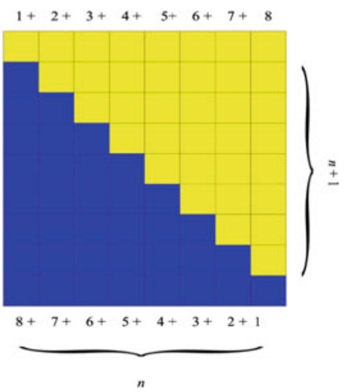
Consider the sum $1+2+3+4+5+6+7+8+9+10$. Write this sum, and the reverse, and add them:

$$\begin{array}{l} 1+2+3+4+5+6+7+8+9+10 \\ \underline{10+9+8+7+6+5+4+3+2+1} \\ 11+11+11+11+11+11+11+11+11+11 = 10 \times 11 \end{array}$$

Because the sum was added to itself, dividing 10×11 by 2 gives the sum.

Argument 2

Consider the sum: $1+2+3+4+5+6+7+8$



$$\Rightarrow 1+2+3+\ldots+n = \frac{n(n+1)}{2}$$

Argument 3

n is either odd or even.

First, consider an odd n, for example 7. Then the sum is $1+2+3+4+5+6+7$. You can rearrange this is to 3 pairs: $1+7, 2+6, 3+5$, all adding up to 8, with the 4 in the middle left out. So the sum is $3 \times 8+4$, or in general $\left(\frac{n-1}{2}\right)(n+1)+\left(\frac{n+1}{2}\right)$, which simplifies to $\left(\frac{n(n+1)}{2}\right)$.

Next, consider an even n, for example 8. Then the sum is $1+2+3+4+5+6+7+8$. You can rearrange this is to 4 pairs: $1+8, 2+7, 3+6, 4+5$ all adding up to 9. So the sum is 4×9, or in general $\left(\left(\frac{n}{2}\right)(n+1)\right)$, which simplifies to $\left(\frac{n(n+1)}{2}\right)$.

At the level of school mathematics, Stylianides (2007b) provides a conceptualization of the meaning of proof that was built taking into account the four major elements of any argument:

> The four elements are the argument's *foundation* (i.e., what constitutes its basis: definitions, axioms, etc.), *formulation* (i.e., how it is developed: as a logical deduction, as a generalization from particular cases, etc.), *representation* (i.e., how it is expressed: using everyday language, algebraically, etc.), and *social dimension* (i.e., how it plays out in the social context of the community wherein it is created). (Stylianides 2007a, p. 2)

In Stylianides' terminology the term "formulation" is ambiguous. Does it refer to the reasoning going on in the mind of the author, or in the mind of the reader, or is it independent of any mind? Our examples illustrate this ambiguity. For instance, as Argument 1's authors we were thinking of 10 as a generic example, and the argument as a proof, as it would work in exactly the same way for any number. But a reader might assume we chose 10 as an example that is sufficiently large to be a typical number, so that if it works for 10 it would probably work for other numbers. In Balacheff's (1988) terminology, the reader sees 10 as a "crucial experiment", not a generic example. Or in Aberdein's (2012) terms we intended our argument to point to a deductive inferential structure, but the reader might not follow that pointer. Thus the argument has two possible "formulations". In one the argument uses logical deduction on a generic example, and we would consider it a proof. In the other, it is a generalization from a particular case that has been chosen to be typical, but is not seen as general. This ambiguity means "formulation" cannot be independent of a mind; the determination of the formulation of an argument depends on a psychological process occurring in a reader (who might be the author of the argument). To be a proof we believe a generic argument must be truly generic, and that depends on a psychological process that might be different for different readers. Something similar can be seen with the other two arguments above.

Social Factors

Even if a generic argument is psychologically a proof, in that it points to a deductive inference structure, it may still not be socially acceptable as a proof. Stylianides (2007a) argues that "the convincing power of an argument is by itself not enough to capture the social dimension of proof in school settings" (p. 12). We believe that what is psychologically convincing and what is socially convincing are mostly different. "A proof becomes a proof after the social act of 'accepting it as a proof'. This is true of mathematics as it is in physics, linguistics, and biology" (Manin 1977, p. 48). And it is also true in classrooms.

> An argument that could count as proof in a classroom community should be accepted as proof by the community – and, thus, it should be convincing to the students – on the basis of socially accepted rules of discourse that are compatible with those of wider society. (Stylianides 2007a, p. 15)

Hence, the arguments above might be differently considered from psychological and social perspectives. It might be a proof in a school community or in ancient Greece, but might not be a proof in the university mathematical community.

Limiting our view of proofs by just taking into account the "fussiness" of an argument would not allow us to acknowledge that proof and proving exist also in social settings, like schools, with their own criteria for proof.

As we noted, at the level of the mathematicians' community the criteria are broadly known by all members belonging to that community. But what happens at school level? Do school students know "a priori" what a generic argument is? Do they really understand this kind of arguments (either those they write or read)? How can a mathematics (school) teacher know his/her students actually understand generic arguments? What would be the classroom criteria for agreeing if and when an argument is a proof? Can those class rules be determined? How?

We believe that one key point to take into account here is clear rules guiding the classroom work in the context of proof, which should also include the case of generic arguments. We return to this in the next section.

Implications for Education: Connecting Psychological and Social Factors

Many authors have pointed out the importance of working with generic examples (e.g., Balacheff 1988; Kempen and Biehler 2015; Malek and Movshovitz-Hadar 2011; Mason and Pimm 1984). In any case, we may say that generic arguments are powerful tools as they can make proof construction accessible to students at any level.

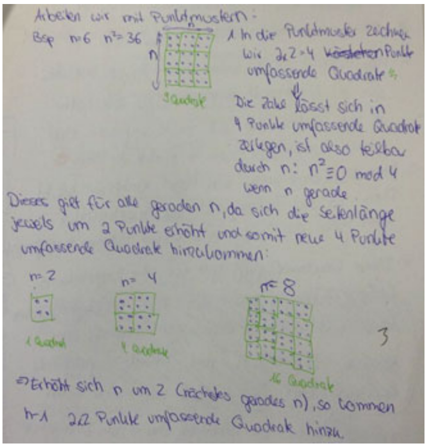
As we have outlined above, determining if a generic argument is a proof or not cannot be done solely on the basis of a characteristic of the text itself, like 'fussiness'. This is the 'absolutist' perspective Stylianides et al. (2016) refer to when discuss perspectives that can be considered in relation to the function(s) of a proof. We adopt what they call a 'subjectivist' perspective, in which psychological processes occurring in readers and the author of the text are considered, and in addition we consider the standards for proof in the community. We suggest an intertwined relationship between the psychological and the social factors as a way to include the use and understanding of generic arguments in classroom settings.

In mathematics teaching, the teacher is setting and the students are learning the classroom standards for proof in part through the acceptance or rejection of arguments. If generic arguments are to be accepted in classrooms (as Yopp et al. 2015, among others, have advocated) then it is important to provide teachers and students with a framework in which to decide if a generic argument is a proof or not. There also needs to be a framework for the social acceptance of proofs that overcomes the limitation that each individual has access only to their own psychological processes.

In the following sections we outline some criteria that could form such a framework. This framework seeks to establish a bridge between the psychological and social factors considered above. This interconnection between these two factors is relevant in the sense that it can promote the kind of explicitness necessary in classrooms when discussing whether or not a generic argument is a proof. And as Selden (2012) notes, "Understanding and constructing such proofs entails a major transition for students but one that is often supported by relatively little explicit instruction" (p. 392). Even though the author refers to proofs in general, we believe that in this context, making this framework explicit in classrooms when working with generic arguments might help, first, students to be more aware of what their considerations are when involved in generic arguments writing, and second, it might also help teachers to have in some way access to the students' psychology.

The Need for Further Examples

Fischbein (1982) observed that once a mathematics statement has been proven, there should be no need for further examples. While this is not categorically true, as examples can have purposes beyond verification (see Lockwood et al. 2012), a significant difference between a generic example and a *specific* example is that additional *specific* examples add nothing to a generic example. A generic example provides a model for the generation of endless specific examples, removing the need to actually produce them. However, multiple specific examples can be important in the formulation of an argument, as a systematic variation of examples can be used to reveal the structure of a generic example (see Fig. 17.1).



The statement being proven is that the square of an even number is always divisible by 4. The text reads "*We work with a dot pattern: E.g.,* $n = 6$ $n^2 = 36$
1st. In the dot pattern we draw $2 \times 2 = 4$ *dots making squares*
⇒ *The number can be divided into squares encompassing 4 dots, and so is* <u>*divisible by n*</u>*:*
$n^2 \equiv 0 \bmod 4$ *if* n *even.*
This is true for all even n*, because as the side length increases by 2 dots, so new squares encompassing 4 dots are added.*
As n *increases by 2 (next even* n*),* $n-1$ *squares encompassing* 2×2 *dots are added.*"

In the underlined portion, we believe that when the student wrote divisible by n she meant divisible by 4.

Fig. 17.1 A generic proof using several examples to show how the structure applies to other cases

In the classroom, the teacher can ask whether the students believe they need more examples to verify a statement, and based on their need for more examples and the use they make of those examples (to provide empirical evidence or to reveal structure) the teacher can determine if the students understood the examples as generic or specific. On the other hand, in written work (for instance in written tests) one might need to have other kind of evidence. We will return to this point later (see *Reconstructing psychology from text*).

Understanding

An important difference between a evidence and an argument with a specific example, or even a set of specific examples, is that a generic argument can be explanatory. It is difficult to define exactly what makes a proof explanatory. Steiner (1978) suggests that

> An explanatory proof makes reference to a characterizing property of an entity or structure mentioned in the theorem, such that from the proof it is evident that the result depends on the property. It must be evident, that is, that if we substitute in the proof a different object of the same domain, the theorem collapses; more, we should be able to see as we vary the object how the theorem changes in response. (p. 143)

When applied to a proof using a generic example, this suggests that for an example to be (psychologically) a generic example, it must be possible to see that as the example varies in some ways the theorem remains true. For example, in Argument 1, the final number in the sum can be larger or smaller, but the numbers must be consecutive and must begin with 1 for the argument to work. The argument can reveal to a reader the properties of the example that are important and in this way they explain the theorem. This means that a reader can answer questions like “Why divide by 2?”, “Why multiply by n + 1?” when given either Argument 1 or Argument 2. Not all arguments are equally explanatory. Argument 3 makes it more difficult to understand the division by 2, because the reason for it is slightly different in the two cases. In contrast, however, a simple example only shows that the formula works for that example, not *why*.

Reconstructing Psychology from Text

In written work, the teacher does not have the immediate opportunity to ask the student questions to determine their thinking. And it is sometimes necessary to decide if a written text is a generic argument that qualifies as a proof without asking further questions (for example when evaluating examinations). In educational settings, it is desirable for both teachers and students to have clear criteria for the evidence that should be included in written work. This evidence should be sufficient to allow the teacher to determine whether the students' attention is on particular/specific examples or a general structure. We suggest that two kinds of evidence should be included in written work:

(1) Evidence of awareness of generality;
(2) Mathematical evidence of reasoning.

Evidence of awareness of generality can be revealed by phrases such as "the same reasoning can be used for the other cases", or "it also applies to the other cases involved", or as in the example seen in Fig. 17.1, the student includes: "this is true for all even *n*". This kind of evidence shows that the author is actually aware that she is working with an argument that is general enough to be valid in all cases involved in her statement. This evidence must be part of the awareness the students should have when presenting a generic argument. The main reason of considering this as relevant evidence is the need to be sure whether or not the students are aware that they are not only dealing with empirical evidence, but that their work shows general structures through the use of their examples. If this is the case, they should include this as part of their written work.

Mathematical evidence of reasoning, reveals the form of the reasoning behind the argument. This kind of evidence mainly points to the mathematical reasons for why the same structure can be extrapolated for other cases from the example(s) given, and it is based not only on the conditions of the problem given but also on the ground knowledge the community shares at that point (the social aspect). For example, in Fig. 17.1, the student is using certain assumptions which seem to be accepted in the context of her class: the square of a number "n" is a square of dots with "n" rows and "n" columns of dots; a number is divisible by 4 if you can make groups of 4 dots without having any dot without grouping it, etc. And based on these assumptions and her data (she is only working with even numbers), she provides the mathematical reasons of why the conclusion holds through the use of her examples.

Both kinds of evidence are relevant when working with generic arguments. One might think of a student including the first kind of evidence in her written work, but if she does not provide the reasons (mathematical evidence of reasoning) that support her "apparent" awareness of generality, then the student's argument could not qualify as a proof in this classroom environment. Or vice versa, if a student works on a general well-structured argument through the use of examples, but if she does not see it as general (she is not aware of this generality), then it is (psychologically) not a proof for that student. In any case, it is a challenge for teachers to determine whether or not an argument based on an example (or a set of examples) is a generic proof without having sufficient evidence of both kinds.

In this context, the following questions can be considered to guide students when writing generic arguments:

(1) Did you state that the argument can be applied to all other cases in discussion? Can it?

(2) Did you describe the reasons (sufficient deductive evidence) behind the generic argument? That is, have you identified the underlying structure in the example or examples, and shown why it occurs in every case?

We think that having in mind these two kinds of evidence, the students can be more aware of what is expected from them when writing a generic argument so it is clearer that they regard it whether or not as a proof, and it might help a reader reconstruct the author's psychology.

These classroom criteria for the kind of evidence that should be provided must be known by the classroom community. They are also part of the social context for generic arguments in the classroom. Indeed, these criteria establish a link between the psychological and social factors since they indicate both what the author finds convincing using deductive reasoning and whether the argument conforms to the classroom conventions.

For example, in Argument 1, it is not clear enough that the author is aware that the argument applies to any natural number. A reader might have this awareness of generality, but it is a mistake to assume an author has the same awareness without having evidence. Argument 1[a] shows a variation of Argument 1 that includes evidence of awareness of generality.

Argument 1[a]

Consider the sum $1+2+3+4+5+6+7+8+9+10$. Write this sum, and the reverse, and add them:

$$\begin{array}{l} 1+2+3+4+5+6+7+8+9+10 \\ \underline{10+9+8+7+6+5+4+3+2+1} \\ 11+11+11+11+11+11+11+11+11+11 = 10 \times 11 \end{array}$$

Because the sum was added to itself, dividing 10×11 by 2 gives the sum.

The same reasoning can be used for any natural number n, and not only for the case of 10.

The addition of the final line provides evidence of awareness, but the form of the reasoning behind the argument is still not clear. Argument 1[b] shows a variation of Argument 1[a] that also includes mathematical evidence that reveals the form of the reasoning behind the argument.

Argument 1[b]

Consider the sum $1+2+3+4+5+6+7+8+9+10$. Write this sum, and the reverse, and add them:

$$\begin{array}{l} 1+2+3+4+5+6+7+8+9+10 \\ \underline{10+9+8+7+6+5+4+3+2+1} \\ 11+11+11+11+11+11+11+11+11+11 = 10 \times 11 \end{array}$$

> Adding down always gives 11 (one more than the last term in the sum) because in the first case we are adding $10 + 1$ (the last term plus 1), and then we are adding a number that is one more (1 becomes 2) to a number that is one less (10 becomes 9). The numbers are consecutive, so the increase in the top row is the same as the decrease in the bottom row. There is one sum adding down for every number in the top row, which is 10 in this case. So we multiply the highest number in the sum (10) by one more than the highest number (11). The product is two times bigger than it should be, because we added $(1 + 2 + \cdots + 9 + 10)$ twice, so to find the real sum we divide the product (10×11) by 2. The same reasoning can be used for any natural number n, and not only for the case of 10.

With the addition of these last lines the status of Argument 1 becomes less ambiguous. As readers trying to reconstruct the psychology of the author of Argument 1, the evidence included in Argument 1[b] gives us a basis to believe that the author was aware of the generality of the argument, and used deductive reasoning to arrive at the conclusion. Thus this argument meets the criteria for a generic argument to be a proof in the classroom. It not only makes the reader aware of the general character of the argument, but also reflects the author's awareness of this generality.

Kempen and Biehler (2015) call the text we added in Argument 1[b] "narrative reasoning" and say such a text should accompany an argument using generic examples in order for it to be considered a proof.

> It is the narrative reasoning that follows the generic examples, which makes a generic proof a valid general argument. So it gets possible to stress the differences between purely empirical examples and valid general arguments. (p. 137)

However, we do not suggest that the *only* way to provide mathematical evidence is with the use of narratives. Some students might feel confident using written words to express their ideas of generality, but others might struggle with linguistic formulations and be better able to use other representations to express the same idea. Argument 1[c] (in Fig. 17.2) shows an alternative way to present Argument 1 [b], without the use of much written language.

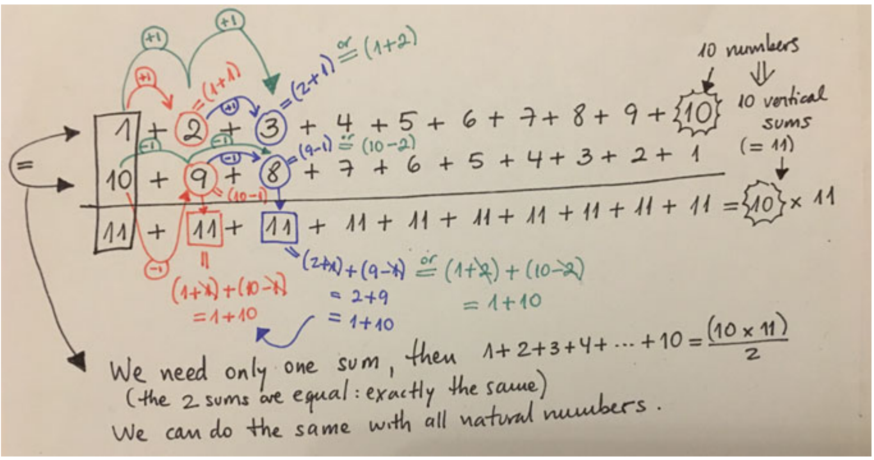


Fig. 17.2 Argument 1[c] in which markings and labels replace most of the verbal narrative

In both cases (Arguments 1[b] and 1[c]) the general structure of the argument has been pointed out, and there is evidence of awareness of generality. Hence these two examples of generic arguments can be considered proofs, according to the criteria suggested for the context of written work in a classroom.

Conclusion

Examples can be used when providing an argument that might qualify as a proof. Depending on whether or not the general structure through the use of examples has been pointed out in an argument (with the use of narratives, or other representations), students can count on a more accessible way to present proofs in classroom settings.

In this article we have argued that the criterion of fussiness is inadequate to decide if generic arguments are proofs. Instead we suggest two other requirements, one psychological and one social. Psychologically, for a generic argument to be a proof it must result in a general deductive reasoning process occurring in the mind of the reader that convinces the reader that there exists a fully deductive inference structure behind the argument. Socially, for a generic argument to be a proof it must conform to the social conventions of the context. In school classrooms or in ancient Greece a generic argument might be acceptable proof. In a university classroom exactly the same argument might not be. Searching for properties of an argument that make it a proof is insufficient because being a proof depends on psychological and social factors independent of the argument. In school classrooms the social conventions are partly based on mathematical criteria, but also on the need in schools for students to convince the teacher that the relevant awareness and reasoning occurred.

Mason and Pimm (1984) raised several questions about generic examples when introducing the concept:

> How can you expose the genericity of an example to someone who sees only its specificity? Apart from stressing and ignoring, and repeating the general statement over and over, how can the necessary act of perception, of seeing the general in the particular, be fostered?
>
> How can you discern the extent of the generality perceived by someone else when looking at a particular example together?
>
> Why do we offer students examples in class, and what are they supposed to make of them? If examples are always examples of something, how can students become aware of that which the examples are supposed to be exemplifying? (pp. 287–288)

We hope that we have addressed in part these questions, and that we have contributed to the ongoing conversation on the role of generic examples in proof and proving.

References

Aberdein, A. (2012). The parallel structure of mathematical reasoning. In A. Aberdein & I. J. Dove (Eds.), *The argument of mathematics* (pp. 351–370). Dordrecht: Springer.

Balacheff, N. (1988). Aspects of proof in pupils' practice of school mathematics. In D. Pimm (Ed.), *Mathematics, teachers and children* (pp. 316–230). London: Hodder and Stoughton.

Epstein, R. L. (2012). Mathematics as the art of abstraction. In A. Aberdein & I. J. Dove (Eds.), *The argument of mathematics* (pp. 249–279). Dordrecht: Springer.

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

Hardy, G. H. (1928). Mathematical proof. *Mind, 38,* 11–25.

Hersh, R. (1993). Proving is convincing and explaining. *Educational Studies in Mathematics, 24* (4), 389–399.

Kempen, L., & Biehler, R. (2015). Pre-service teachers' perceptions of generic proofs in elementary number theory. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the 9th Congress of the European Society for Research in Mathematics Education* (pp. 135–141). Prague, Czech Republic.

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

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

Malek, A., & Movshovitz-Hadar, N. (2011). The effect of using transparent pseudo-proofs in linear algebra. *Research in Mathematics Education, 13*(1), 33–57.

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

Mason, J., & Pimm, D. (1984). Generic examples: Seeing the general in the particular. *Educational Studies in Mathematics, 15*(3), 277–289.

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

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

Selden, A. (2012). Transitions and proof and proving at tertiary level. In G. Hanna & M. de Villiers (Eds.), *Proof and proving in mathematics education: The 19th ICMI study* (pp. 391–422). Heidelberg: Springer Science + Business Media.

Steiner, M. (1978). Mathematical explanation. *Philosophical Studies, 34,* 135–151.

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

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

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

Yopp, D., Ely, R., & Johnson-Leung, J. (2015). Generic example proving criteria for all. *For the Learning of Mathematics, 35*(3), 8–13.

Chapter 18
Systematic Exploration of Examples as Proof: Analysis with Four Theoretical Frameworks

Orly Buchbinder

Abstract This chapter offers a multi-layered analysis of one specific category of students' example-based reasoning, which has received little attention in research literature so far: systematic exploration of examples. It involves dividing a conjecture's domain into disjoint sub-domains and testing a single example in each sub-domain. I apply four theoretical frameworks to analyze student data: The Mathematical-logical framework for the interplay between examples and proof, Proof schemes framework, Transfer-in-pieces framework, and the Theory of instructional situations. Taken together, these frameworks allow to examine the data from mathematical, cognitive and social perspectives, thus broadening and deepening the insights into students thinking about the relationship between examples and proving. Implications for teaching and learning of proof in school mathematics are discussed.

Keywords Example-based reasoning · Proof by cases · Proof-schemes
Transfer in pieces · Instructional situations · Multiple theoretical perspectives

Introduction

Students' reliance on empirical evidence for proving general statements is a widespread and well-documented phenomenon (e.g., Balacheff 1988; Healy and Hoyles 2000). In theorizing about its nature, researchers proposed several useful constructs such as proof schemes (Harel and Sowder 1998) and example-based reasoning (Healy and Hoyles 2000). Researchers have also distinguished between example *types* and example *uses* in proving by students (Ellis et al. 2013) and by mathematicians (Lockwood et al. 2016). Other studies have focused on how students understand the interplay between examples and proving (Buchbinder 2010) and on the relationship between example generation and proof production (Alcock

O. Buchbinder (✉)
University of New Hampshire, Durham, USA
e-mail: orly.buchbinder@unh.edu

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_18

and Weber 2010; Sandefur et al. 2013). These studies have shown that example generation can support proof production, when examples can illuminate the underlying structure of the proof, or definitions of key mathematical concepts in the conjecture. Similarly, counterexamples can provide insights into why a particular conjecture is false, and serve as a vehicle to refine the conjecture.

Despite the reports on productive uses of examples in proving, prevailing evidence shows students rely on supportive examples for drawing general conclusions. This chapter contributes to this body of knowledge by taking a closer look at one specific category of students' example-based reasoning—*systematic exploration of examples*. It involves dividing a domain of a statement into disjoint subdomains, and testing a random example in each subdomain to determine the truth-value of the statement. This category was identified in a study that examined high-school students' conceptions of the roles of examples in proving (Buchbinder 2010). This category has been chosen for closer analysis not because of its commonality, but because even though the overall reasoning was incorrect, students' arguments seem to involve quite sophisticated mathematical thinking. Hence, deeper analysis of responses in this category might shed light on the reasoning processes underlying students' thinking, inform our understanding of students' conceptions of proving and suggest potential mediating solutions.

I start by presenting the task, followed by two responses produced by students, which illustrate *systematic exploration of examples*. Next I will interpret the data using four theoretical frameworks and discuss how these different types of analysis illuminate and complement each other.

The Task and the Data

The following task was given to 6 pairs of high-attaining 10th grade students (7 girls and 5 boys) from two Israeli high-schools. All students had prior experience with proofs in the context of high-school geometry. In the Israeli curriculum, high-attaining students study Euclidean geometry with emphasis on proof in grades 9 and 10. Geometry is studied concurrently with algebra, which emphasizes developing proficiency with algebraic techniques. Thus, although students were proficient in algebra, their experience with proofs in algebra was limited.

The students volunteered to participate in the study, which consisted of a series of task-based interviews, conducted separately with each pair of students by the author of this chapter. The mathematical content of the tasks was taken from the regular middle or high-school curriculum, although the specific tasks were unfamiliar to students in both content and structure. The tasks were given in a paper and pencil form, with no time limitations. Students were allowed to ask clarifying questions about the wording of the tasks. All interviews were videotaped and transcribed for analysis.

During the third interview session, the students were presented with the following task:

The task: A fraction between two fractions
A student took two fractions $\frac{1}{2}$ and $\frac{3}{4}$, and added the two numerators and the two denominators in the following way: $\frac{1+3}{2+4} = \frac{4}{6} = \frac{2}{3}$. The student noticed that the resulting fraction $\frac{2}{3}$ is between the two original ones: $\frac{1}{2} < \frac{2}{3} < \frac{3}{4}$. Is this a coincidence?

The hypothetical student's observation leads to a conjecture, which can be written: *For any two fractions* $\frac{a}{b} < \frac{c}{d}$*, it is true that:* $\frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}$. The conjecture is true, and is known as the *mediant property*. Several versions of this task have been used in prior research, for example, Bishop (2001) discussed it in the context of teacher decision making in the midst of instructional activity. Rowland and Zazkis (2013) as well as Stylianides and Stylianides (2010) discuss multiple approaches for proving the mediant property[1] as well as applications of versions of this task for teacher education.

In the context of the study in which the data were collected, the mediant property was introduced within *Is this a coincidence?* type of task (Buchbinder and Zaslavsky 2011). The task consists of the description of steps taken by a hypothetical student, and an observation he/she makes following these steps. The distinctive design feature of this type of task is that the conjecture is not stated explicitly; its formulation is left to students, and may vary based on their interpretation of a single example presented in the task. For example, one pair of students noticed that the numerators and denominators of the two fractions $\frac{1}{2}$ and $\frac{3}{4}$, are four consecutive natural numbers, and assumed this to be a set of relevant fractions. Hence, they formulated and proved that a median property holds for such fractions. In the two data cases, discussed in this chapter, students took the domain of the task to be all fractions.

The question accompanying the description and the observation—is this a coincidence?—implies the need to test whether the observation made by the student is unique to the single example tested, or an instance of a general rule, which would then require a proof. To assist the students to interpret the question—is this a coincidence?—as intended by the researcher, the task included two additional prompts. One, required students to write down a conjecture as they interpret it from the task; and second, describe what they think is needed in order to prove or disprove the conjecture, that is to determine whether the observation made by the student is a coincidence, or not. Successful completion of the task might involve testing a few numeric examples to convince oneself that the statement is true, followed by an algebraic proof to show that the observed property is "not a

[1]See Nelsen (1993) for several elegant proofs without words of the mediant property.

coincidence" (Buchbinder and Zaslavsky 2011). Note that in this particular task, *A fraction between two fractions*, examples can provide hardly any insight into the underlying structure of the proof. In order to prove the assumed conjecture, one needs to change the mode of inquiry from testing examples to creating and manipulating an algebraic representation of the mediant property.

Such algebraic proof is within the mathematical competence of secondary students; indeed, three out of six student pairs produced it. One student pair found what they thought was a counterexample, but in fact was a calculation error; two remaining pairs of students justified the conjecture through an approach which I term *systematic exploration of examples*, and it is illustrated below.

Illustrative Student Response 1

Neta and Ronit[2] divided the domain of all fractions into two subdomains: proper and improper fractions, and produced, what they thought was a proof of the mediant property by examining one randomly chosen pair of fractions in each subdomain, and additional pair of fractions with one fraction chosen from each of the two sub-domains. As their final answer the students produced the following justification, they wrote:

Both fractions smaller then 1:	$\frac{1+1}{4+2} = \frac{1}{3} \to \frac{1}{4} < \frac{1}{3} < \frac{1}{2}$ true.
One fraction <1, another fraction >1:	$\frac{2+6}{3+3} = \frac{4}{3} \to \frac{2}{3} < \frac{4}{3} < \frac{6}{3}$ true.
Both fractions bigger then 1:	$\frac{5+6}{2+4} = \frac{11}{6} \to \frac{6}{4} < \frac{11}{6} < \frac{5}{2}$ true.

In order to prove that when you add two numerators and two denominators the resulting fraction is always between the original ones, you need to prove it with other cases that will fit the rule. Like here, we tried different ones and we proved it that it is not a coincidence.

Note that Neta and Ronit did not consider the case where both fractions are equal to 1, or the case in which one fraction is equal to one, and the other is not, since they interpreted the task as being about two non-equal fractions. The next example illustrates a similar line of reasoning, but with less conventional partition of the domain of conjecture.

Illustrative Student Response 2

Tami and Natalie noticed that in fractions $\frac{1}{2}$ and $\frac{3}{4}$ the denominator is greater than the numerator by 1, and asserted that the conjecture is true for all pairs of fractions

[2]All names are pseudonyms.

which have equal difference between the numerator and the denominator, that is, pairs of fractions: $\frac{a}{b}$ and $\frac{c}{d}$, such that $b - a = d - c$. Therefore, the students focused on two sub-domains: in their words, "fractions that have the same difference between the denominator and the numerator," such as $\frac{4}{7}$ and $\frac{6}{9}\,(7 - 4 = 9 - 6)$ and fractions that do not have this property, for example: $\frac{5}{12}$ and $\frac{3}{4}\,(12 - 5 \neq 4 - 3)$. Tami and Natalie proved the conjecture by checking one example in each sub-domain by saying:

Tami: This is what we need to prove this [conjecture]. Cases. Here we took fractions with the same difference and we got: $\frac{4}{7} < \frac{4+6}{7+9} < \frac{6}{9}$ which is true. And here we took fractions where the difference is not the same, and we got the correct answer $\frac{5}{12} < \frac{5+3}{12+4} < \frac{3}{4}$. This means that this is always true.

When pressured to explain their reasoning the following interaction occurred:

Interviewer: I am trying to understand…. In all other tasks, you were constantly saying things like: "I want to use variables, I want α and x, the numbers bother me." And here ….

Tami: Because this is a completely different thing! In Geometry, you can play with angles, you have 180º and so on. And here, you can't play with it. In algebra, it's better to prove with cases: smaller, equal to, inequalities and things like that. You can substitute numbers because it's easier than with x-es; x-es are for general cases when you want to prove that something is true. But here, for each special case…Like we did here. Let's say, the difference is the same—it's true; the difference is not the same—it's also true. So, this is what you need—substitute numbers. Not x-es.

Interviewer: Is this something that you would write on a test?

Tami: Yes, yes. Maybe, I would add something with x-es as well. But if I write $\frac{x}{x+1}$ it's the same as writing $\frac{3}{4}$. Here the difference is the same, and here it's the same. So, what does it matter if I substitute x or a number? It will be the same. So, I prefer numbers. It's much easier.

Interviewer: Easier…and do you think it is equivalent?

Tami: Yes.

Towards Interpreting Students' Responses

Given the setting of the task-based interviews as a part of research study, the amount of time students devoted to the task (approximately 20 min), and their level of engagement with it, it is unlikely that students' responses are the result of carelessness or time constraints; nor are they a product of insufficient algebraic knowledge. Students' use of *systematic exploration of examples* as a proving strategy was a conscious choice. Their verbal and written justifications reflect confidence in the appropriateness of their solutions.

Other salient features of students' answers include referring to their examples as *cases*, rather than examples; the use of words *proof* and *proof by cases*; and multiple references to inequalities. In mathematics, a proof by cases involves dividing a domain of a statement into exhaustive cases, and proving a conjecture for each case, or defining a set of subdomains that span the domain, and proving the conjecture for each subdomain. For example, certain statements in elementary number theory are proved by considering the domains of even and odd numbers separately. It is possible that the two pairs of students thought they were applying proof by cases by simply checking specific examples in each subdomain, instead of proving the conjecture holds there.

Another interesting aspect of students' responses is their multiple reference to inequalities. To find a solution set of an inequality of the form $y < ax + b$, one first graphs the equation $y = ax + b$ which divides the plain into three regions: the two half plains on each side of the line itself (in case of a strict inequality the line itself is not a part of the solution set). Then, a common procedure is to test a single random point in each region to determine the solution set.

The point-testing strategy for solving inequalities is introduced in many secondary algebra textbooks with minimal or insufficient justification. Boero and Bazzini (2004) note that this practice is common in many countries and might lead to "a "trivialisation" of the subject, resulting in a sequence of routine procedures, which are not easy for students to understand, interpret and control." (p. 140). But there are some notable exceptions when this strategy is presented with careful attention to its mathematical validity. For example, in CME Project's Algebra 1 textbook (Cuoco et al. 2013), a solution set to a linear inequality in two variables is described as a collection of rays, where each ray is a solution to a one-variable inequality for a fixed value of x. From this explanation, a point-testing strategy is developed as a simplification of examining multiple rays. But when the focus of instruction shifts from justifying the procedure to performing it fluently, it is not clear to what extent students remain aware of theoretical grounds underlying and restricting its applicability.

Solving linear, quadratic and rational inequalities is a part of Israel's grade 10 curriculum, and point-testing is one of the common ways students are taught to solve them. Hence, it is possible that students associated the task *A fraction between two fractions* with solving inequalities, and applied some version of the point-testing procedure learned in the context of inequalities to that proving task.

In the following I apply four theoretical frameworks to analyze the data: The Mathematical-logical framework for example-proof interplay (Buchbinder and Zaslavsky 2009), Proof schemes framework (Harel and Sowder 1998, 2007), Transfer-in-pieces (Wagner 2006, 2010), and the Theory of instructional situations (Herbst and Chazan 2012). My goal of choosing these frameworks was not to be exhaustive about possible ways to analyze the data, clearly, alternative or additional frameworks might be suggested further. These particular frameworks were chosen because they allow to analyze mathematical responses of students working in pairs, not just individuals. In addition, the frameworks represent three different perspectives: mathematical (The Mathematical-logical framework for example-proof interplay), cognitive (Proof schemes framework and Transfer-in-pieces framework), and social (the Theory of instructional situations). As such, each framework provides an additional layer of insight into student thinking about the relationship between examples and proving. In this work, I follow Cobb's (2007) suggestion to view different theoretical perspectives as sources for ideas which can illuminate different aspects of students' mathematical activity.

Note, that since a comprehensive description of four theoretical frameworks is not possible within the scope of a single chapter, the frameworks are only briefly outlined. The reader is referred to the references for the extension.

Interpreting the Data with Four Theoretical Frameworks

Mathematical-Logical Framework for the Interplay Between Examples and Proving

Buchbinder and Zaslavsky (2009) proposed a mathematical-logical framework that describes four types of examples: supporting (or confirming), contradicting (or non-confirming), and irrelevant (type 1 and type 2). The status of each type of example with respect to proving or disproving a mathematical statement depends on the type of statement: universal or existential. This status can be one of the following: sufficient (for proving or disproving); insufficient; indicating impossibility to prove or disprove (e.g., confirming example indicates that an existential statement cannot be disproved); or not enough information (e.g., supporting example indicates uncertainty whether a universal statement is true or false). The framework describes the status of all types of examples with respect to the type of statement.

Students' responses can be mapped on to different aspects of the framework and analyzed in terms of their alignment, or the lack of thereof, with conventional mathematics, represented by the framework. Instances of alignment, such as a student disproving a general statement upon discovering a counterexample, are considered indicators of understanding of the relevant aspect of the framework; e.g., a status of contradicting example in refuting a universal statement. Responses misaligned with conventional mathematical reasoning are considered *non-normative* rather than as indicative of misunderstanding. This apparent asymmetry in the interpretation of mathematically correct and incorrect student work stems from the notion that erroneous student responses can mask other elements of conceptual understanding or alternative mathematical thinking (Buchbinder 2010; Ron et al. 2010).

Since the object of the analysis is a mathematical response, the framework can be applied to analyze responses of individual students as well as students working in pairs or even small groups. Analyzing a set of responses to a collection of carefully chosen mathematical tasks allows to diagnose specific area, or areas, in which students' responses are aligned or misaligned with conventional mathematics, and reveal strengths and weaknesses in students' understanding[3] of the examples-proof interplay.

Applying this framework to the two cases presented above, shows that both student pairs rely on the limited number of supporting examples for proving a general conjecture. This inappropriate use of inductive inference is misaligned with conventional mathematical reasoning, and therefore is considered a non-normative response. As such, it can potentially indicate students' problematic conception: In accepting the illegitimacy of generalizing from a small number of supportive examples. However, it can also reflect an alternative mathematical reasoning, or the possible miss-application of otherwise correct mathematical ideas. Examining the data from students' perspective required application of additional analytic tools, as described below.

Proof-Schemes Framework

Students' responses can be interpreted using Harel and Sowder's (1998) proof schemes framework, which helps to identify sources of conviction for a particular person or a community. Harel and Sowder (1998, 2007) define a proof scheme as a set of processes that an individual, or a community, employs to convince themselves or others whether a certain assertion is true or false. The inclusion of a community in the definition of a proof scheme makes this framework appropriate for analyzing justifications produced by pairs of students. The taxonomy of proof schemes consists of three classes, each of which has several sub-classes. In the *external conviction* proof schemes the source of conviction resides within an external authority, such as a book or a teacher, (the authoritarian proof scheme), the appearance of an argument (the ritual proof scheme), or the presence of symbolic manipulations (the non-referential symbolic proof scheme). In the *empirical* proof scheme the source of conviction is either empirical evidence (the inductive proof scheme) or perceptual clues (the perceptual proof scheme). The *deductive* proof scheme is characterized by increased reliance on generality, operational thought and logical inference.

In the responses above students utilize a combination of strategically and randomly selected examples to prove a general conjecture. These data can be interpreted as evidence of the inductive proof scheme, meaning that what constitutes the source of conviction and an acceptable mode of justification for the students is empirical evidence, i.e., supportive examples. However, certain elements of students' responses can be interpreted as evidence of other proof schemes: a transformational proof scheme, which is a sub-category of a deductive proof-scheme, or

[3]According to the framework, understanding is operationalized as consistent application of inferences that are aligned with conventional mathematical knowledge.

an external conviction proof scheme, specifically, ritual and authoritarian proof schemes.

The data suggest that students seem to attribute greater generality to their examples than might initially appear. For instance, Tami's claim of equivalence between $\frac{x}{x+1}$ and $\frac{3}{4}$, and both student pairs' use of the word *cases* rather than examples, suggest that students viewed their examples as representative of some general entities, or as generic examples (Mason and Pimm 1984; Leron and Zaslavsky 2013). Harel and Sowder (2007) place a proof by generic example within the transformational proof scheme. Although there is not enough evidence to claim that students indeed viewed their examples as generic, the data show that, rather than relying on the examples per se, both student pairs justified their reasoning by virtue of the *systematic process* involved in choosing and testing these examples—the process which they call "proof by cases." Possibly, both pairs of students were under impression that they applied some version of proof by cases in their response. Viewed from this perspective, students' proofs could be expressions of the transformational proof scheme, in the sense that students were perusing a goal of proving a conjecture for all cases by applying, what they thought, was valid logical inference.

On the other hand, students' response can be an instance of application of point-testing strategy, which they came to associate with solving of inequalities. Viewed from this perspective, the proofs produced by students are indicative of the ritual proof scheme and the authoritarian proof scheme, in the sense that students might have developed a ritual for how proofs involving inequalities should look like, instigated by the authority of a teacher of a textbook. The legitimacy of testing a single or small number of examples in the context of solving certain types of problems in school algebra, especially when supported by textbooks, can interfere with students' developing conceptions of mathematical proof as a general argument. Students who hold an external conviction proof scheme receive implicit, unintentional support for reliance on supporting examples.

Transfer-in-Pieces Framework

Scholarship on the transfer of knowledge offers an additional lens for examining students' responses. Building on Smith et al.'s (1993) epistemology of knowledge-in-pieces Wagner's (2006, 2010) approach of transfer-in-pieces maintains that mathematical knowledge is dynamic and highly contextualized. The way an individual interprets a particular mathematical situation is the result of complex interactions between one's prior knowledge and the situational context. Similarly, an individual interprets two mathematical situations as "similar" or "different" based on the available knowledge resources and contextual *cues* of the situation. Wagner (2010) explains that "to say that resources are *cued* is to say that they are actively available and accessible for use, likely serving to provide the language and explanatory or inferential concepts by which the individual structures the situation"

(p. 452). It is also important to note, that the transfer-in-pieces framework recognizes that the social situation in which the problem solving takes place can influence "which ideas have high cueing priority" (ibid.). Thus, application of the transfer-in-pieces framework seems appropriate for analyzing proving activity of pairs of students.

The data above suggests that students' knowledge of both solving of inequalities and proving by cases is fragmentized—although students acquired correct mathematical ideas in these areas, they seem to apply them inappropriately. This is not surprising, since intermediate and partial understanding are natural stages in the process of growth of knowledge. Moreover, an individual who has been sensibly using a certain concept in some contexts, may not have developed sufficient knowledge resources that enable him or her to attend to and interpret available information that calls for applying this concept in another contextual situation. Wagner terms such set of knowledge resources a *concept projection*, and asserts that the individual who have not have developed sufficient concept projection may "still struggle or be entirely unable to make use of the concept in a new or unfamiliar situation" (Wagner 2010, p. 451). Wagner also suggests that once particular knowledge resources are cued, they guide how an individual contracts the problem situation in the way that it appears valid to the individual and affords him or her to function and interact meaningfully with the problem and with other learners.

Applying the transfer-in-pieces framework to student responses, there is some evidence to suggest that students have developed correct, but partial mathematical knowledge of a concept of proof by cases, and of the procedure for solving linear or rational inequalities. Yet, they have not developed sufficient concept projection that will allow them to apply these concepts and procedures sensibly in novel situations. It is possible than, that certain contextual features of *A fraction between two fractions* proof-task, such as inequality sign, and algebraic fractions, cued students' knowledge of solving inequalities. Once these concepts have been cued, they triggered the use of language and inferential resources appropriate in that context, such as point-testing approach, rather than solution strategy which would be more appropriate in the context of proving, including proving by cases.

Instructional Situations Framework

The perspectives discussed above can be further augmented by applying Herbst and Chazan's (2012) theory of instructional situations. An instructional situation is a system of tacit expectations and implicit norms that students and teachers develop around a particular mathematical content or task type. These norms specify the types of solutions students are expected to produce that would be deemed acceptable by teachers. For example, Buchbinder et al. (2015) discuss how in the instructional situation for solving equations teachers perceived correct mathematical solutions that deviate from the standard, canonical method for solving equations (e.g., dividing all terms of an equation by a common factor as a first step) as a-typical and less appropriate than canonical solutions.

In their research Herbst and Chazan explore the construct of instructional situations as a way to study the practice of teaching. But if the norms of instructional situations develop as mutual expectations of teachers and students, it is reasonable to assume that students would hold the same, or similar, types of norms of certain instructional situations. I hypothesize that in the instructional situation for solving inequalities, point-testing is a type of solution that would be considered normative and acceptable by both students and teachers. However, in the instructional situation for doing proofs, testing a few examples would be considered inappropriate and rightfully rejected by a teacher.

The construct of instructional situations provides a useful lens to analyze students' responses. The students in the study had prior experiences with proving in geometry, and with solving inequalities in algebra. During the interview, Tami made a clear distinction between acceptable modes of justification in geometry and algebra by saying that they are "completely different." Although she was aware that "x-es are for general cases when you want to prove that something is true," she seemed to interpret the task as being about solving inequalities, saying: "In algebra, it's better to prove with cases: smaller, equal to, inequalities and things like that." Tami, therefore, recognized proving in geometry and solving algebraic inequalities as two distinct instructional situations, with different expectations for solving strategies. Hence if the students interpreted the given proof task as being about solving inequalities, rather than a situation for proving a general statement, they might have acted according to expectations they developed for what constitutes an appropriate solution in the instructional situation for solving inequalities, instead of applying solution strategies that are appropriate for the instructional situation for proving.

Discussion

This chapter examined a special category of example-based reasoning: *systematic exploration of examples*. Previous studies that explored students' uses of examples in proving reported on such strategies as using either random examples, or examples that follow particular patters, for instance, common examples, unusual examples, or boundary cases (Ellis et al. 2013). Bell (1976) has identified a category of example-based reasoning called "empirical-systematic", in which students, attempted to exhaust all possible sets of cases, in a task that has a finite set of such cases. Balacheff (1988) used the term "crucial experiment", to describe proving strategy that relies on carefully selected set of examples. The types of student arguments discussed in this chapter constitute an additional category of example-based reasoning, which is a combination of the two previously discussed in literature: use of random examples and use of systematic examples. Its uniqueness stems from the perceived generality students associated with it, considering it a “proof by cases.” Indeed, students’ solutions bear some resemblance to this form of mathematical proof, particularly in partitioning the domain of conjecture into disjoint sub-domains.

The four theoretical frameworks used to interpret the data should not be viewed as competing, but rather as complementing each other, with each framework providing conceptual tools for examining the data from a slightly different perspective. Buchbinder and Zaslavsky’s mathematical-logical framework takes the conventional mathematical knowledge as a starting point. Its application allowed to diagnose the particular area in which students’ responses are misaligned with conventional mathematical knowledge—understanding the role and limitations of using supporting examples in proving. The mathematical-logical framework categorizes such students’ responses as non-normative, or simply misaligned with conventional mathematics, as opposed to erroneous, thus, signaling a need to further examine students’ reasoning strategies and look for potential rationality in their answers. However, since this framework does not conceptualize the perceived generality of examples by students, additional analytic tools were used to gain further insights into students’ thinking.

The data were analyzed using Harel and Sowder’s proof schemes framework, to identify modes of reasoning that students find convincing and acceptable for proving. A proof scheme is a “collective cognitive characteristic of the proofs one produces” (Harel 2007, p. 265), hence it is it is not possible to determine or classify participating students’ proof schemes on account of a single proof they produced. However, since each proof is “a product of mental act of proving, characterized by a certain proof scheme” (Harel 2007, p. 266), the analysis carried above aimed to identify characteristics of certain proof schemes within the specific proofs produced by the students.

This analysis suggested that although students’ proofs can be manifestations of an inductive proof scheme, alternative proof schemes can be involved. In particular, students might have been following a point testing procedure introduced by a teacher and/or a textbook, which would be consistent with a ritual or authoritative proof scheme. Alternatively, students might be under impression that they are implying proof by cases, viewing their examples as general “cases” representative of a relevant sub-domain—an approach that could be interpreted as manifestation of the transformational proof scheme.

Students’ references to solving inequalities and proof by cases were further analyzed using Wagner’s transfer-in-pieces framework and by applying Herbst and Chazan’s theory of instructional situations. Wagner’s transfer-in-pieces framework explored cognitive mechanisms underlying students’ responses, such as contextual cues associated with inequalities and algebraic fractions. These cues might have triggered the relevant cognitive structures which provided inferential and linguistic resources for students to draw on in their solutions. The application of the transfer-in-pieces framework allowed to examine the data from students’ perspective, focusing on the cognitive mechanisms and the sources of rationality underlying their responses. It helped to identify elements of correct mathematical knowledge and theorize about how they came into play in students’ solution approaches. The theory of instructional situations (Herbst and Chazan 2012) adds a socio-cultural dimension to the analysis by locating the sources of students' responses in tacitly held social expectations for what constitutes appropriate and acceptable solution of certain types of tasks in a mathematics classroom. Collectively, the four theoretical frameworks provided a variety of tools that illuminated different facets of the data.

Cobb (2007) cautions that, although applying multiple theoretical frameworks can be beneficial, researchers should be mindful of potential conflicts between them, and if needed, apply appropriate adaptations. One way in which this issue was addressed in this study, is by choosing theoretical frameworks that are rooted within the field of mathematics education, as opposed to using theories from other fields [e.g., see Leron and Hazzan (2009) for an example of analysis that uses cognitive and evolutionary psychology alongside mathematics education theories]. In addition, two of the frameworks specifically address students' conceptions of proving, and three of them originated in the research on learning, except for the theory of instructional situations, which originated from the research in teaching. Table 18.1 summarizes the theoretical perspective used in this study, their origins and what they afforded in terms of the data analysis.

The kind of analysis carried out in this chapter can be potentially extended to analyzing other instances of students' use of examples in proving. Bringing together multiple theoretical perspectives can deepen and broaden the analysis of students' proofs, and enhance the field's understanding of students' conceptions related to using examples in proving.

Implications for Education

The Mathematical-logical framework for the interplay between examples and proving (Buchbinder and Zaslavsky 2009) characterizes *systematic exploration of examples* as misaligned with general strategies for proving universal statements. Hence, it is important that students develop understanding of the differences between *systematic exploration of examples* and other valid modes of reasoning for testing and proving conjectures, including proof by cases. The analysis, carried out in this chapter, can potentially inform design of instructional tasks and practices that capitalize on correct mathematical ideas embedded in *systematic exploration of examples* to compare and contrast it with valid proving strategies. Similarly, analyzing the problematic aspects of this example-based strategy can serve as a "springboard for inquiry", in Borasi's (1994) terms.

For example, when teaching the topic of solving inequalities, teachers can emphasize the specificity of applicability of point-testing by discussing with students why is this strategy valid for this particular purpose. It can then be further contrasted with proof by cases and with using examples for exploring, rather than proving, conjectures. Careful design of the appropriate instructional activities can draw on the analysis afforded by application of Wagner's (2006, 2010) theory of

Table 18.1 Summary of the four theoretical frameworks used in this study

The framework	Originated in research on	Focus of the analysis in this study	Affordances of the analysis
Mathematical-logical framework for the interplay between examples and proving (Buchbinder and Zaslavsky 2009)	Students' conceptions of proving	How students' responses compare to conventional mathematical knowledge	Diagnose the area in which students' responses are misaligned with conventional mathematical knowledge
Proof schemes framework (Harel and Sowder 1998, 2007)	Students' conceptions of proving	Identifying manifestations of proof schemes in students' proofs	Identify possible manifestations of authoritative and/or transformational proof schemes, as alternatives to inductive proof scheme
Transfer-in-pieces (Wagner 2006, 2010)	Students' transfer of knowledge within mathematics	Salient features of the task, reflected in students' proofs, that cued certain language and inferential resources	Suggested cognitive mechanisms underlying students' responses. Identify elements of correct mathematical knowledge in students' responses
Theory of instructional situations (Herbst and Chazan 2012)	Teachers' tacit norms and expectations. Was adapted based on a hypothesis that these norms would be shared by students	Social norms and expectations that are specific to certain instructional situations	Located the source of students' rationality in tacitly held, content specific social expectations

Transfer-in-pieces, by recognizing how the salient features of mathematical tasks, such as solving inequalities and proving general conjectures trigger students' language and inferential resources. By clarifying the similarities and differences among the various ways of using examples in proving teachers can help students to develop concept projections that will allow students to use examples sensibly and appropriately in proving tasks.

The critical role of teacher in this process is particularly emphasized by the theory of the instructional situations (Herbst and Chazan 2012). Since social norms are held tacitly, it is possible that teachers are unaware of the conflicting messages regarding the roles of examples that are conveyed to students in the instructional situation of solving inequalities, or other instructional situations that do not specifically focus on proving. Promoting change in this direction would require supporting teachers in developing such awareness and creating new instructional situations, in which the role of examples in proving is emphasized and contrasted with other uses of examples in mathematics.

References

Alcock, L., & Weber, K. (2010). Referential and syntactic approaches to proving: Case studies from a transition-to-proof course. In F. Hitt, D. Holton, & P. Thompson (Eds.), *Research in collegiate mathematics education VII* (pp. 93–114). Washington: AMS.

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

Bell, A. W. (1976). A study of pupils' proof-explanations in mathematical situations. *Educational Studies in Mathematics, 7*(1–2), 23–40.

Bishop, A. J. (2001). Educating student teachers about values in mathematics education. In F.-L. Lin & T. J. Cooney (Eds.), *Making sense of mathematics teacher education* (pp. 233–246). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Boero, P., & Bazzini, L. (2004). Inequalities in mathematics education: The need for complementary perspectives. In M. J. Høines & A. B. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 139–143). Bergen, Norway.

Borasi, R. (1994). Capitalizing on errors as "springboards for inquiry": A teaching experiment. *Journal for Research in Mathematics Education, 25*(2), 166–208.

Buchbinder, O. (2010). *The role of examples in establishing the validity of universal and existential mathematical statements*. Unpublished Doctoral dissertation. Technion, Haifa (in Hebrew).

Buchbinder, O., Chazan, D., & Fleming, E. (2015). Insights into the school mathematics tradition from solving linear equations. *For the Learning of Mathematics, 35*(2), 1–8.

Buchbinder, O., & Zaslavsky, O. (2009). A framework for understanding the status of examples in establishing the validity of mathematical statements. In M. Tzekaki, M. Kaldrimidou, & C. Sakonidis (Eds.), *Proceedings of the 33rd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 225–232). Thessaloniki, Greece.

Buchbinder, O., & Zaslavsky, O. (2011). Is this a coincidence? The role of examples in fostering a need for proof. *ZDM—The International Journal of Mathematics Education, 43*(2), 269–281.

Cobb, P. (2007). Putting philosophy to work. Coping with multiple theoretical perspectives. In F. K. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 3–38). Charlotte, NC: Information Age.

Cuoco, A., et al. (2013). *CME algebra 1 common core*. Boston, MA: Pearson.

Ellis, A. E., Lockwood, E., Williams, C. C. W., Dogan, M. F., & Knuth, E. (2013). Middle school students' example use in conjecture exploration and justification. In L. R. Van Zoest, J. J. Lo, & J. L. Kratky (Eds.), *Proceedings of the 34th Annual Meeting of the North American Chapter of the Psychology of Mathematics Education* (pp. 135–142). Kalamazoo, MI: Western Michigan University.

Harel, G. (2007). The DNR system as a conceptual framework for curriculum development and instruction. In R. Lesh, J. Kaput, E., & Hamilton (Eds.), *Foundations for the future in mathematics education* (pp. 263–280). Mahwah, NJ: Lawrence Erlbaum Associates.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Research in collegiate mathematics education III* (pp. 234–283). Washington, DC: Mathematical Association of America.

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

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education, 31*(4), 396–428.

Herbst, P., & Chazan, D. (2012). On the instructional triangle and sources of justification for actions in mathematics teaching. *ZDM—The International Journal of Mathematics Education, 44*(5), 601–612.

Leron, U., & Hazzan, O. (2009). Intuitive vs. analytical thinking: four perspectives. *Educational Studies in Mathematics, 71*(3), 263–278.

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

Lockwood, E., Ellis, A. B., & Lynch, A. G. (2016). Mathematicians' example-related activity when exploring and proving conjectures. *International Journal of Research in Undergraduate Mathematics Education, 2*(2), 165–196.

Mason, J., & Pimm, D. (1984). Generic examples: Seeing the general through the particular. *Educational Studies in Mathematics, 15*(3), 277–289.

Nelsen, R. B. (1993). *Proofs without words: Exercises in visual thinking*. Washington, DC: Mathematical Association of America.

Ron, G., Dreyfus, T., & Hershkowitz, R. (2010). Partially correct constructs illuminate students' inconsistent answers. *Educational Studies in Mathematics, 75*(1), 65–87.

Rowland, T., & Zazkis, R. (2013). Contingency in the mathematics classroom: Opportunities taken and opportunities missed. *Canadian Journal of Science, Mathematics and Technology Education, 13*(2), 137–153.

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

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

Stylianides, G. J., & Stylianides, A. J. (2010). Mathematics for teaching: A form of applied mathematics. *Teaching and Teacher Education, 26*(2), 161–172.

Wagner, J. E. (2006). Transfer in pieces. *Cognition and Instruction, 24*(1), 1–71.

Wagner, J. E. (2010). A transfer-in-pieces consideration of the perception of structure in the transfer of learning. *The Journal of the Learning Sciences, 19*(4), 443–479.

Chapter 19
Using Examples of Unsuccessful Arguments to Facilitate Students' Reflection on Their Processes of Proving

Yosuke Tsujiyama and Koki Yui

Abstract Proving is an essential component in mathematical activities, but a difficult one for many students. We assume that one reason for this might be that Unsuccessful arguments unsuccessful arguments made during the process of planning a proof do not appear in the completed proof, and therefore students cannot see how those arguments influenced the proof. If students could reflect on such arguments, they would be able to learn about proving and effective ways to derive a proof. Previous studies have provided worked examples showing successful ways of deriving a proof to enhance students' understanding of proving. However, such examples do not include unsuccessful arguments. This chapter examines how examples of unsuccessful arguments can facilitate students' reflection on their process of planning a proof by designing, implementing, and analyzing an eighth-grade geometry lesson. It was found that an example of unsuccessful arguments enabled the students to comprehend why the unsuccessful arguments failed and why the successful ones worked.

Keywords Proof and proving · Planning a proof · Reflection · Argumentation
Unsuccessful arguments · Worked-out example

Introduction

Students are expected to learn various aspects of proof and proving, which are central to their experience of school mathematics. For example, they learn what a proof is, how to establish a proof, and why they need a proof (Hanna and Barbeau 2008; Harel and Sowder 2007; Heinze et al. 2008). Among these, this study focuses on students' learning about proving itself, that is, it examines students' processes of

Y. Tsujiyama (✉)
Chiba University, Chiba, Japan
e-mail: tsujiyama@chiba-u.jp

K. Yui
Shinshu University, Matsumoto, Japan

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_19

proving and the ways in which planning and constructing a proof take place (Boero 1999; Polya 1957/2004).

Planning a proof is considered an argumentative process in which students make plausible and probable arguments, although they are not sure if such arguments can be successfully used in the proof. Previous studies have analyzed the argumentation during the processes of conjecturing and revealed factors that influence the success or failure of the proof (Garuti et al. 1996; Pedemonte 2007). They focus on a learning context in which students make a conjecture and then prove it, and they analyze the relationship between the argumentation and the completed proof. In contrast, this study focuses on the argumentative processes involved in proving a given statement and finding new statements to be proved next. In particular, this chapter focuses on the processes involved in planning a proof, reflecting on the planning processes, including on the unsuccessful arguments that do not appear in the completed proof, and tries to apply the findings of the previous studies to this context.

Although the arguments made during the process of planning a proof may be unsuccessful, they may help students in learning about proving. Particularly, reflection on unsuccessful arguments will promote students' understanding of the reasons why successful ones worked, while the others did not. This understanding can help them develop effective ways of planning a proof and finding new problems (Tsujiyama 2011).

Unsuccessful arguments are usually removed during the process of constructing a proof and do not appear in the completed proof. Therefore, without teachers' guidance, students will not be able to review how those arguments influenced the proof. This leads to the idea of a "heuristic worked-out example" (Reiss et al. 2008). This idea is to be applied to the case of unsuccessful arguments, to help students learn about the process of planning a proof. Thus, this chapter aims to examine how examples of unsuccessful arguments facilitate students' reflection on their process of planning a proof.

Theoretical Perspective

Reflection on Processes of Proving that Include Unsuccessful Arguments

Proving a given statement involves two phases: planning and constructing a proof. Planning a proof is seeking how to connect premises and conclusions of a statement deductively, and constructing a proof is organizing a connected sequence of deduction between the premises and conclusions (Tsujiyama 2012). Processes of proving do not happen in a straightforward manner, but moves back and forth even in mathematicians' studies (Boero 1999; Polya 1957/2004). As students are not sure what the expected proof is like, they may make errors during planning a proof and may have to modify them during constructing the proof (Heinze et al. 2008). In such cases, they need to seek and organize deductive connections again. We focus on this tentative nature of proving.

From the argumentative perspective (Toulmin 1958/2003), proving is seen as making plausible and probable arguments in planning a proof and examining those arguments deductively in constructing the proof. If the arguments made are unsuccessful on examination, students have to seek other arguments again in planning a proof (Tsujiyama 2012). Therefore, the simplest process that includes unsuccessful arguments comprise the following phases: (1) making arguments in planning a proof; (2) examining the arguments and removing unsuccessful ones in constructing the proof; (3) seeking alternative arguments in planning the proof; and (4) refining the alternative arguments and constructing the completed proof. We propose that if students experience such processes and reflect on them, they will be able to understand why their initial arguments failed and why alternative arguments worked. Moreover, they will learn more effective ways of planning a proof.

Several studies in mathematical philosophy discuss aspects of a proof as a product, although not about the processes of proving, similar to this reflection on unsuccessful arguments. This aspect is called the “explanatory” function of proof (Hanna 1989). Roughly, an explanatory proof is one that has a power to answer not only that a statement is true but also why the statement is true. Weber and Verhoeven (2002) consider not merely a single theorem and its proof but also related theorems and their proofs (or disproofs). They suggest that a couple of explanatory proofs (one of them can be a disproof) answer a why-question of the form, “Why do mathematical objects of class X have property Q, but not property Q’?” (Weber and Verhoeven 2002, p. 304). Thus, they consider that it is important to answer the why-question by contrasting a theorem and its proof with related theorems and their proofs (or disproofs).

Weber and Verhoeven consider the function of proofs and the why-question to be answered as the characterizing property of the theorem to be proved. In contrast, in this chapter, we focus on arguments made during the process of planning a proof and the question to be answered as “Why did or did not the arguments work?” However, the two viewpoints can be compared as follows. The former clarifies why a statement is true by contrasting it with why the related statements are true (or false). The latter provides the reason for the working of a successful argument by contrasting it with the reason for the failure of an unsuccessful one. For example, when students face a proof problem of the form “Given X, show Y,” they can find several reasonable intermediary premises by backward reasoning from Y (Heinze et al. 2008). The students may select P1 as one of the premises, go on the bridging process to X, and successfully obtain a proof. They may also select P2, find it impossible to bridge X and Y by P2, and fail to obtain a proof. In this situation, both arguments, P1 implies Y and P2 implies Y, are correct. On the other hand, X implies P1 is correct while X implies P2 is incorrect. Our assumption is that reflecting on these arguments will make students realize why P1 works and why P2 does not, and this realization will lead an effective way of selecting successful intermediary premises in planning a proof.

Use of Examples of Unsuccessful Arguments

With regard to phases (1)–(4) mentioned previously, the central issue is that arguments made and removed in (1) and (2) do not appear in the completed proof. Therefore, however carefully students check the proof, they cannot see how those arguments influenced it. Thus, we need to consider teachers' instructions to facilitate students' reflection. This leads to presenting specific type of examples similar to the students' activities in (1) and (2). In this respect, we consider "heuristic worked-out examples" (Reiss et al. 2008).

A heuristic worked-out example comprise a problem, its detailed solution, and heuristic strategies that guide the problem-solving. Reiss et al. (2008) examined studies in cognitive science and focused on (traditional) worked-out examples, which consist of a problem and its detailed solution. According to them, when students are required to solve problems with regular instruction, they often lack an understanding of the underlying mathematical principles and solve problems with strategies that are fundamentally shallow. This may lead to a solution but not enhance students' understanding since such strategies occupy cognitive resources in working memory. On the contrast, (traditional) worked-out examples support students in gaining understanding. Reiss et al. applied this idea for effective instruction in learning of proving and considered more process-oriented examples. Thus, they included heuristic strategies and introduced "heuristic" worked-out examples in their work (Reiss et al. 2008, pp. 457–458).

Their idea is based on the assumption that with "regular instruction", students often do not gain mathematical understanding by simply solving a given problem. This is even more significant when learning about planning a proof of a given statement since its process does not usually appear in the completed proof. Thus, we apply this idea for learning about planning a proof.

In a heuristic worked-out example, imaginary students engage in proving activities that mirror the work of expert mathematicians. They begin with finding a problem, examine it, consider mathematical properties related to it, figure out ideas for a proof, and successfully construct a proof. Teachers present this successful example before or during students' proving activities.

Since our focus is to facilitate students' reflection on their past processes, we present an example after the students' proving activities. In addition, since we expect students to reflect not only on the successful arguments that appear in the proof, but also on the unsuccessful ones that have been removed, we provide examples of unsuccessful arguments. Moreover, we focus on unsuccessful arguments that can be obtained through backward reasoning from the conclusion of a given statement since such arguments are effective for deriving a proof and can be used to produce new problems after proving. Presenting such examples, we intend to encourage students to reflect on their processes of planning a proof based on the examples and become aware of why their initial arguments failed and why the modified ones were successful.

Therefore, we define an example of unsuccessful arguments as an example of imaginary students' proving that consists of a plan obtainable through backward reasoning from the conclusion of a given statement, and the resulting arguments that cannot complete the proof. To facilitate students' reflection on their processes of proving, we present the example after students have finished their proving activities. We then lead them to compare the example and the completed proof and examine the processes, especially focusing on unsuccessful arguments that were made in (1) but removed in (2) and therefore, unclear in the produced proof through (3) and (4).

Related to utilization of such examples that have explanatory power for students, Peled and Zaslavsky (1997) proposed the notion of counter-examples that explain. Our notion of an example of unsuccessful arguments does not offer an explanation of why a statement is or is not true, as the counter-examples do, but possibly explains why an argument does or does not work. We examine this through designing, implementing, and analyzing a lesson.

Methods

We collaboratively designed a geometry lesson using an example of unsuccessful arguments. The first author developed the task and its aim theoretically, which involved the expected flow of students' activities and teacher's instruction. The authors then together refined the design of the lesson. The second author implemented the lesson (50 min) with 37 eighth-graders at a junior high school in Japan. The first author observed the lesson and analyzed how the example facilitated students' reflection on their processes of planning a proof.

First, we chose and arranged a suitable problem (Fig. 19.1) such that students might make unsuccessful arguments through backward reasoning naturally based on their prior experience, but could still utilize such arguments to modify their plan toward a proof and to find new statements to be proved next. The students were required to create a pair of triangles to which AB and DC respectively belong (e.g. △ABC and △CDA by drawing diagonal AC). We intended the students to understand why they need diagonal AC (Shimizu 1994) through reflecting on their processes.

Before coming to eighth-grade, students had studied about the various properties of parallelograms in empirical ways. In the eighth-grade, they learnt the terms

Problem

Prove that in a quadrilateral ABCD, if AB || DC and AD || BC, then AB = DC.

Fig. 19.1 Main problem

"proof" and "definition", and they set up several definitions of geometrical objects and fundamental assumptions (e.g., conditions for congruent triangles and properties of parallel lines and angles). They learnt that they could prove statements based on these definitions and assumptions and also how to devise a plan of a proof of statements related to triangles. Especially to prove that two segments are equal in length, they learnt to focus on "finding" a pair of triangles to which the two segments belong respectively as their sides. However, they now need to "create" such triangles by themselves. Thus, the problem was the first opportunity for them not only to prove the properties of parallelograms but also to create such triangles in proving geometrical properties.

In this particular context, it is natural that some students may draw both diagonals (AC and BD) creating another pair of triangles AOB and COD, which seem congruent and to which AB and CD respectively belong. However, they cannot show that triangles AOB and COD are congruent (noted as $\triangle AOB \equiv \triangle COD$ in Japan) since they do not have any pair of equal sides. This assumption is plausible in the above-mentioned context in which students have not proved any properties of parallelograms, including the fact that diagonals in a parallelogram intersect at their midpoints. We focus on this idea as an example of unsuccessful arguments.

We prepared three worksheets to facilitate students' reflection after constructing a proof, which also enabled our analysis. In sum, the students were:

- given the statement that they were supposed to prove;
- given and required to fill in Worksheet 1 about their plan and resulting proof;
- shown two students' results that completed the proof on the blackboard;
- shown the example of unsuccessful arguments on the blackboard;
- given and required to fill in Worksheet 2 about their reflection on their processes of proving; and
- given and required to fill in Worksheet 3 about their comparison between successful proofs and unsuccessful arguments.

First, the students were given the main problem (Fig. 19.1). Then they devised their plan, wrote it down, and constructed a proof in Worksheet 1. After two students who successfully reached correct proofs presented their results on the blackboard, the teacher showed the students the example of unsuccessful arguments. The students compared and examined the similarities and differences of these results, and then filled in Worksheet 2, which included the following three blank columns: (i) what you were troubled by, (ii) how you settled it, and (iii) what you learned today and what you want to do next. Worksheet 3 included the example of unsuccessful arguments and a blank column: (iv) what is the significance of drawing AC or BD? However, due to time constraints, the teacher could not hand out Worksheet 3 during the lesson. He handed it out at the beginning of the next lesson and the students filled in column (iv).

We used six video cameras and collected data from the entire classroom instruction and three individual students' proving activities. First, we selected one of the three students who had proposed a plan and arguments similar to the ones that we had prepared in the unsuccessful example. We analyzed the student's processes of proving through the transcripts, video record, and Worksheet 1. We then analyzed the student's reflections through Worksheets 2 and 3.

Results

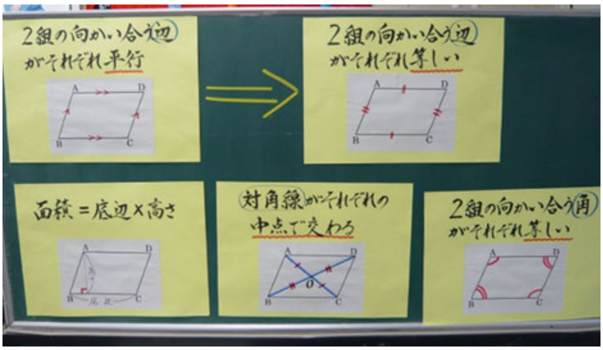
Processes of Proving

During the lesson, the teacher first asked the students about the empirical knowledge they had gained about parallelograms at the elementary schools; the students answered that opposite sides are parallel (Fig. 19.2a); opposite sides are equal in length (Fig. 19.2b); the area is calculated by base times vertical height (Fig. 19.2c); the diagonals intersect at their midpoint (Fig. 19.2d); and opposite angles have equal measures (Fig. 19.2e). The teacher put up the following images on the blackboard and explained that the problem was to prove (b) based on (a). The students then began solving the problem individually.

The students tackled the problem by marking several segments and angles of equal size or parallel lines and gradually grasped the difficulty in finding a pair of triangles to which AB and DC respectively belong. Many of them came up with an idea of creating such triangles by drawing additional lines, while others could not proceed and seemed to be stuck. At least five students drew both AC and BD but could not complete their proof.

Mizu, one of the students (all names are pseudonyms), initially tried to prove it by drawing both diagonals AC and BD. When 7.5 min had passed, she heard the teacher say to another student that they were not allowed to use (c), (d), and (e). She was surprised and cried out "What?" The teacher came over and talked to her.

(a) (b)



(c) (d) (e)

Fig. 19.2 Properties of parallelograms

101	Teacher:	What happened? What trouble do you have?
102	Mizu:	Triangles … I tried to make triangles, as we always prove by triangles …
103	Teacher:	Oh! You made them! Are this and this [pointing at △AOB and △COD in Mizu's worksheet] congruent? Then, what? Are you not able to prove that? Why?

Mizu looked at the blackboard and pointed at (d).

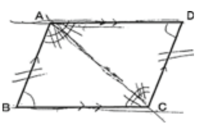
104	Mizu:	I made [△AOB and △COD], made uh, but the diagram became similar to that [pointing at (d)], similar to that where "the diagonals intersect at their midpoint"

Mizu appeared to clarify her thinking through her conversation with the teacher. At that point, she nearly gave up.

113	Teacher:	What? Are you not able to prove?
114	Mizu:	Maybe not.
115	Teacher:	Why?
116	Mizu:	Well, that center one [pointing at (d)].
117	Teacher:	You mean, you cannot prove without using that?
118	Mizu:	I cannot prove without using that.

Then, Etsu, the student sitting next to Mizu, joined the conversation and advised her to change the plan. The teacher urged Mizu to consider Etsu's advice, and left to talk to other students.

Mizu followed Etsu's advice and changed her approach. She erased BD and other markings, except for AC, and then marked the angles ∠DAC, ∠CAB, ∠DCA, ∠ACB. The teacher then came back to Mizu and confirmed with her the properties she had already learnt before. Then, she reviewed her notebook. After a while, she found out and wrote down "alternate interior angles of parallel lines are equal" in Worksheet 1 and cried out. She had an idea and wrote it down on Worksheet 1: "to prove AB = CD, it is fine if I can show congruency of △ABC and △CDA that contain AB and CD," and continued to write a proof. She finally completed the proof by showing △ABC ≡ △CDA. During the process, she erased all the words and almost all the markings in the diagram that she had written down before she changed the plan, including △AOB and △COD (Fig. 19.3).



Draw additional line from A to C.
Alternate interior angles of parallel lines are equal.
Common sides.
Conditions for congruent triangles

Fig. 19.3 Mizu's diagram and plan

Reflection on the Process of Proving

After the teacher confirmed that all the students had written down a plan and proof in their own way, he asked two students to present their results. One showed △ABC ≡ △CDA and the other △ABD ≡ △CDB. After the presentations, the teacher asked all the students, "Why was it necessary to draw AC?" (169 Teacher) The teacher called on some students to answer and gradually elicited the ideas that they had used in planning, that is, creating a pair of triangles to which AB and DC respectively belong. Then the teacher presented the example of unsuccessful arguments and outlined Mikio's (an imaginary student in the example) plan and proof (Fig. 19.4). The teacher emphasized that Mikio was embarrassed since he could not show △AOB ≡ △COD.

The teacher confirmed with the students that the plan in the example was similar to the successful ones presented by the two students with regard to creating a pair of triangles to which AB and DC respectively belong. Then the teacher raised a question, "Why [did Mikio] get stuck?" (218 Teacher) and asked if anybody had attempted a similar approach. One of the students answered saying, "In the hypothesis, uh, side, each side, uh, we cannot show equality of sides, so we cannot prove" (219 Tobi). Then, Mizu answered thus:

232	Mizu:	We can use only these two, uh, in the above [pointing at (a) and (b) on the blackboard], and we cannot use the three below [i.e. (c), (d), and (e)]

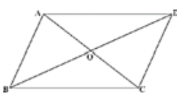
At the end of the lesson, the students wrote down their reflection on Worksheet 2, and four students presented their reflections.

Function of the Example of Unsuccessful Arguments

Mizu filled in columns (ii) and (iii) in Worksheet 2 that she found Etsu's advice to change her viewpoint very helpful. She described what troubled her in column (i) thus:

Mikio's (imaginary student) plan

To derive AB=CD, I will show the congruence of △AOB and △COD to which AB and DC respectively belong.



Mikio's proof

I draw AC and BD and make the triangles AOB and COD to which AB and CD respectively belong. Then I show △AOB ≡△COD to derive the conclusion AB=CD.

In △AOB and △COD, since AB || DC and alternate angles of parallel lines have equal measures, ∠BAO = ∠DCO (1), ∠ABO = ∠CDO (2)

Fig. 19.4 The example of unsuccessful arguments

> Since I was not allowed to use the property [of parallelograms] that the diagonals intersect at their midpoints, I could not connect to the conclusion.

This reflection can be considered as a partial explanation of why her initial idea failed: she was not allowed to use unproven property (d); in other words, it did not cover the possibility that $\triangle AOB \equiv \triangle COD$ can be shown in another way. Moreover, this reflection did not seem to be enabled by the example, since she had stated similar points several times (e.g., 104, 118 Mizu) during the conversation with the teacher before the example was presented.

At the beginning of the next lesson, the teacher handed out Worksheet 3 and asked the question, "What is the significance of drawing AC or BD?" Then, Mizu reviewed the example further, comparing it to her own plan and proof, and wrote down an additional reflection in column (iv):

> Since we have a common side [in $\triangle ABC$ and $\triangle CDA$], we can show the congruency of the triangles. In the case of Mikio, we have two diagonals but no common side.

This description indicates a clearer reason why $\triangle AOB \equiv \triangle COD$ failed, by specifying the absence of a common side. It also explains why the modified argument worked, which is due to the presence of a common side. Therefore, Mizu distinguished between the unsuccessful and successful arguments from the perspective of the absence and presence of a common side. This reflection was facilitated by the example, as she clearly stated, "in the case of Mikio."

This perspective is important didactically and mathematically from two points. One is that when students try to prove a geometrical property by using conditions for congruent triangles, it is effective to find or create triangles that not only seem congruent but also have at least one pair of equal sides. The other is that this distinction can lead sophisticated understanding of the notion of congruency in contrast to similarity.

Thus, the example enabled Mizu to compare successful and unsuccessful arguments and to explain the reason why the successful one worked and the other failed.

Concluding Remarks

In this episode, the example of unsuccessful arguments enabled Mizu to reflect on her process of proving and to clarify the reason why the initial arguments failed and why the modified one worked. This result implies the possibility of using examples of unsuccessful arguments in learning about proving. Particularly, we applied the idea of heuristic worked-out examples, which were introduced by Reiss et al. (2008). However, the approach was different on two counts. First, we presented the example after the students finished their proving activities to facilitate reflection on their processes of proving. Second, we focused on the example of unsuccessful arguments, since we expected students to reflect on not only successful arguments that appear in the completed proof, but also unsuccessful ones that were removed during the processes of proving. These two characteristics are in contrast to the heuristic worked-out examples, which were presented to students before or during the processes of proving and typically used successful arguments.

This difference resulted in a different effect on the students' learning. According to Reiss et al. (2008), heuristic worked-out examples were effective especially for low-achieving students. In our case, Mizu was in trouble for a while, but managed to change her plan and complete a correct proof. This fact implies that we can consider using examples of unsuccessful arguments to facilitate students' reflection on their processes of proving even for high-achieving students.

Further possibilities in using examples of unsuccessful arguments are found in the next lesson. Mizu continued thinking about the unproved property (d) of parallelograms after the lesson was over. She finished it and reported the result in the next lesson, stating that she had proved $\triangle AOB \equiv \triangle COD$. She also found that she could use $\triangle AOB \equiv \triangle COD$ to prove AO = CO and BO = DO, which had troubled her in the previous lesson. This shows the potential of the examples in promoting such problem-posing and problem-solving processes.

In the example of unsuccessful arguments in the lesson, although Mikio's idea and Mizu's initial idea are slightly different, they are also similar in that they both focused on creating a pair of triangles to which AB and DC respectively belong. However, Mikio's arguments failed due to the absence of a pair of equal sides, but Mizu's initial idea failed since she was not allowed to use the unproven property (d). Nevertheless, Mizu not only reflected on her own process concerning the use of (d) but also distinguished between the unsuccessful and successful arguments clearly referring to Mikio's case. Here we see that the example functioned as facilitator to make Mizu understand why the successful argument worked and the other did not. However, we cannot pinpoint how exactly the example facilitated this understanding. We need more detailed analysis and further empirical studies focusing on this process.

Related to the treatment of the unproven property (d) and the definition of parallelograms, Nana, the other student sitting next to Mizu, wrote in column (iii) as follows: "If the premise [of the statement] is not that [AB and DC are] parallel but [that] the diagonals intersect at [their] midpoints, we will be able to use the triangles that Mizu focused on, but conversely not be able to use the triangles that we focused on ([we can] not [use that] alternate-interior angles in parallel lines [have equal measures])." In the lesson, Nana had successfully completed the proof before the conversation of the teacher and Mizu (line 101–118), and listened into the conversation. From our data, it is not clear whether Nana came up with this idea by comparing her own proof and Mizu's initial idea, or by comparing three ideas, including that of Mikio. In the former case, Mizu simply presenting her unsuccessful attempt would have been enough for Nana to gain the above-mentioned understanding of the sophisticated meaning of parallelograms. In the latter case, Nana could have got clearer idea by Mikio's example since Mikio's proof stated more clearly than Mizu's about what could be shown if $\triangle AOB \equiv \triangle COD$ held.

Although Nana did not express this mathematically important idea to others, and it was not taken up in the next lesson, her idea had the potential to lead to the local organization of properties of parallelograms (Freudenthal 1971), refinement of undetermined objects to become under-determined or determined through an examination on their diagrams (Nets 1998), and further inquiry with the refinement of its definition (Borasi 1992). More detailed considerations are needed, especially on the treatment of errors, as examples of unsuccessful arguments for such genuine mathematical inquiries.

The implications of this study are inclined more toward the process-oriented direction of research on proof and proving. Hanna and Barbeau point out that recent studies on proof and proving "seem to have dealt primarily with the logical aspects of proof and with the problems encountered in having students follow deductive arguments" (Hanna and Barbeau 2008, p. 347) emphasizing the educational value of processes of proving over the role of proof as a product. If unsuccessful arguments made during processes of proving have various functions similar to that of proofs, students will have more opportunities to experience the productive aspects of proof and proving.

Acknowledgements We are grateful to Keith Jones (University of Southampton) and the internal reviewers of this book for their critical and constructive comments on the earlier version of this chapter. We also thank Guershon Harel (University of California, San Diego) and Andreas J. Stylianides (University of Cambridge) for their thoughtful supports for this issue. This study was supported by JSPS KAKENHI Grant Number 26780504, 26282039, 16H02068, 17K14033.

References

Boero, P. (1999). Argumentation and mathematical proof: A complex, productive, unavoidable relationship in mathematics and mathematics education. Resource document. *International Newsletter on the Teaching and Learning of Mathematical Proof, 7/8*. http://www.lettredelapreuve.org/OldPreuve/Newsletter/990708Theme/990708ThemeUK.html. Accessed May 1, 2015.

Borasi, R. (1992). *Learning mathematics through inquiry*. Portsmouth, NH: Heinemann.

Freudenthal, H. (1971). Geometry between the devil and the deep sea. *Educational Studies in Mathematics, 3,* 413–435.

Garuti, R., Boero, P., Lemut, L., & Mariotti, M. A. (1996). Challenging the traditional school approach to theorems: A hypothesis about the cognitive unity of theorems. In L. Puig & A. Gutierrez (Eds.), *Proceedings of the 20th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 113–120). Valencia, Spain: PME.

Hanna, G. (1989). More than formal proof. *For the Learning of Mathematics, 9,* 20–23.

Hanna, G., & Barbeau, E. (2008). Proofs as bearers of mathematical knowledge. *ZDM The International Journal on Mathematics Education, 40*(3), 345–353.

Harel, G., & Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 805–842). Reston, VA: National Council of Teachers of Mathematics.

Heinze, A., Cheng, Y.-H., Ufer, S., Lin, F.-L., & Reiss, K. (2008). Strategies to foster students' competencies in constructing multi-steps geometric proofs: Teaching experiments in Taiwan and Germany. *ZDM The International Journal on Mathematics Education, 40*(3), 443–453.

Nets, R. (1998). Greek mathematical diagrams: Their use and their meaning. *For the Learning of Mathematics, 18*(3), 33–39.

Pedemonte, B. (2007). How can the relationship between argumentation and proof be analysed? *Educational Studies in Mathematics, 66*(1), 23–41.

Peled, I., & Zaslavsky, O. (1997). Counter-examples that (only) prove and counter-examples that (also) explain. *Focus on Learning Problems in Mathematics, 19*(3), 49–61.

Polya, G. (2004). *How to solve it: A new aspect of mathematical method* (2nd ed). New Jersey, NJ: Princeton University Press (Original work published 1957).

Reiss, K., Heinze, A., Renkl, A., & Groβ, C. (2008). Reasoning and proof in geometry: Effects of a learning environment based on heuristic worked-out examples. *ZDM The International Journal on Mathematics Education, 40*(3), 455–467.

Shimizu, S. (1994). Demonstration. In S. Shimizu (Ed.), *CRECER Vol. 6 Geometry and demonstration* (pp. 204–236). Tokyo: Nichibun [in Japanese].

Toulmin, S. E. (2003). *The uses of argument (updated edition)*. Cambridge: Cambridge University Press (Original work published 1958).

Tsujiyama, Y. (2011). On the role of looking back at proving processes in school mathematics: Focusing on argumentation. In M. Pytlak, T. Rowland, & E. Swoboda (Eds.), *Proceedings of the 7th Congress of the European Society for Research in Mathematics Education* (pp. 161–171). Rzeszów, Poland: ERME.

Tsujiyama, Y. (2012). Characterizing how to examine probable arguments in constructing a proof in school mathematics: From an argumentative perspective. *Journal of Science Education in Japan, 36*(4), 340–355 [in Japanese].

Weber, E., & Verhoeven, L. (2002). Explanatory proofs in mathematics. *Logique & Analyse, 179–180,* 299–307.

Chapter 20
Genericity, Conviction, and Conventions: Examples that Prove and Examples that Don't Prove

Orit Zaslavsky

Abstract The four chapters in this section address issues related to the use of examples in proof and proving. Several questions arise from reading these chapters. I structure this chapter around some of these questions: First questions related to the nature of (mathematical) examples and their sources, then questions related to generic proving, including the subjective nature of generic proof, different levels of genericity, and how students may view generic arguments. I conclude with some observations regarding rigor and evidence.

Keywords Example-based reasoning · Generic examples · Generic proof
Conviction · Rigor

The four chapters in this section complement each other and together offer a rich perspective on the nature and roles of examples with respect to proof, in different contexts. Buchbinder examines how high school students use examples in the course of forming and verifying a conjecture, related to algebra and number properties (the median theorem); Tsujiyama and Yui examine the process of learning to construct a geometric proof in middle school, with the added experience of discussing an unsuccessful argument; Kempen examines meta aspects related to pre-service mathematics teachers' conceptions of proofs, some of which rely on examples; and Reid and Vallejo Vargas, focus on generic proofs in the context of school mathematics, from a more theoretical point of view.

O. Zaslavsky (✉)
New York University, New York, USA
e-mail: orit.zaslavsky@nyu.edu

O. Zaslavsky
Technion Haifa, Haifa, Israel

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3_20

The Many Facets of Examples

What Constitutes an Example?

The term *example* appears in the literature in various meanings. According to Zaslavsky (2014), an example is an object for which one can answer the question: "What is this an example of?". This view resonates with other approaches to defining examples, such as Watson and Mason's (2005), who describe examples as "objects which satisfy certain conditions" (p. 378), Alcock and Weber's (2010), who maintain that an example is a mathematical object satisfying the definition of some concept, or Mill's (2014), who considers an example a mathematical object that is "a specific, concrete representative of a class of mathematical objects, where the class is defined by a set of criteria" (p. 107).

Goldenberg and Mason (2008) go beyond specifying the conditions that an example should satisfy and maintain that it is situated within a person's understanding and serves as means for making contact with abstract ideas. Furthermore, a mathematical entity is considered to be an example only when the person perceives it to be an instance of a phenomenon, property, class, or idea: "The fundamental construct is the act of *seeing something as an example of some 'thing'*" (ibid, p. 184, emphasis in the original).

The above implies that an example does not stand alone. There is a subjective element that depends on the context and the person using the example, and is connected to both *intention* and *attention*. For example, the number 36 can be viewed as an example of an even number (because it can be represented as 2×18), a multiple of 3 (because it can be represented as 3×12), a composite number (because it can be represented of a product of numbers other than 1 and 36), a perfect square (because it can be represented as 6^2), and so on. Each representation satisfies a defining property and may be used to draw attention to a certain structure. Hence, in the context of example-use, one needs to articulate what the example stands for, as inevitably an example carries some attributes that are intended to be exemplified and others that are irrelevant. Skemp (1987) refers to the irrelevant features of an example as 'noise', while Rissland Michener (1991) suggests that "one can view an example as a set of facts or features viewed through a certain lens" (p. 190). The tendency to attend to irrelevant features of an example may explain the possible disconnect between a concept image and a concept definition (Vinner 1983; Vinner and Dreyfus 1989). According to Vinner, a wrong concept image held by students, e.g., of an isosceles triangle, might be a result of the set of specific examples that they have been exposed to. Thus, if students only saw examples of isosceles triangles that have a horizontal basis, they may attend to this irrelevant feature and consider it critical, and as a result not identify isosceles triangles that do not have a horizontal basis.

The notion of an example can be extended beyond examples of well-defined mathematical concepts. Along this line, the chapters of Tsujiyama and Yui, Reid and Vallejo Vargas, and Kempen look at examples of arguments or proofs as objects of inquiry and try to characterize and evaluate their status and potential merits. Arguments and proofs can be seen as meta-concepts, according to Zaslavsky and Shir (2005), who studied students' and teachers' conceptions of another meta-concept—a mathematical definition. Tsujiyama and Yui focus on the merits of examining examples of imaginary students' arguments that are 'unsuccessful' for a particular proof of a certain statement as a means to facilitate learning to construct a proof. They are inspired by Reiss's et al. idea (2008) of heuristic worked-out example. By analogy, this can be seen as the use of examples and non-examples in the course of constructing a concept (Hershkowitz 1990; Hershkowitz and Vinner 1983). In both cases, the inclusion of examples that 'don't work' aims at highlighting critical features of the relevant concept or meta-concept (proof). Somewhat related, Reid and Vallejo Vargas as well as Kempen look at examples of 'proofs', some based on examples and some more deductive, some rather complete and some that call for additional parts, and question their status. Like in the case of more straightforward mathematical examples, and even more so, there is a strong element of subjectivity in the ways in which these authors view their examples and the properties they attribute to them.

Given the subjective aspect of perceiving an example, a critical question that arises from the chapters has to do with the evidence (or lack of evidence) we have for interpreting students' example-uses. More specifically, how can we know what exactly students are attending to in an example, what they are noticing about it, how they are treating it, what sense are they making of it, etc.? Buchbinder addresses this issue by offering four different yet complementing theoretical frameworks through which she examines students' example-use. Reid and Vallejo Vargas suggest criteria for the evidence that should be included (in students' written work).

What Constitutes a Generic Example?

As Reid and Vallejo Vargas note, Mason and Pimm (1984) made a further distinction and coined the term *generic example*, for an example that for the person using it conveys the general through the particular.

For example, the absolute-value function $f(x) = |x|$ is often used as a generic example of a function that is continuous everywhere but not differentiable everywhere. The issue of representation plays a critical role in generic examples, as the representation can either help or impede the structure or general behaviour that is seen through it. In the case of the absolute-value function, its *graph* may convey to some people the essence of why this is the case, without having to go into symbolic manipulations. Similarly, the cubic function $f(x) = x^3$ is often considered a generic example of an odd function. These examples are what Rissland Michener (1978) considers *model* examples, that capture the general features of a concept (or a result).

In the context of learning more advanced mathematical concepts, Harel and Tall (1991) suggest the use of a specific example that “is seen by the teacher as a representative of the abstract idea” (p. 40). They term this a generic example, and note that students may abstract wrong properties from such an example. Interestingly, they look at a generic example from the point of view of the teacher, as do Movshovitz-Hadar (1988) and Leron and Zaslavsky (2013). From a designer’s point of view, Movshovitz-Hadar (1988) maintains that a generic example should be “large enough to be considered a non-specific representative of the general case, yet small enough to serve as a concrete example” (ibid, p. 17). Leron and Zaslavsky suggest that “size” be replaced by a measure of the *complexity* of the example. For example, when illustrating the *procedure* for finding all factors of a natural number, by listing systematically all its factorizations as a product of two factors, the complexity is measured by the number of factors, not by the magnitude of the number. For this purpose, 169 is less generic than 36 or 42, since the former is too special, having only 3 factors. We can use 42 as a generic example as follows:

$$\begin{array}{c} 1 \times 42 \\ 2 \times 21 \\ 3 \times 14 \\ 6 \times 7 \end{array}$$

~~7 × 6~~

One can see through this example (perhaps with some guidance) the general structure of the procedure, that is, starting with the smallest factor (1) on the left and increasing it without skipping any factor, until it begins to repeat itself, in a reversed order. To better ‘cover’ the general case, it would be worthwhile examining also a perfect square, say 36, as follows:

$$\begin{array}{r} 1 \times 36 \\ 2 \times 18 \\ 3 \times 12 \\ 4 \times 9 \\ 6 \times 6 \end{array}$$

This could be an opportunity to point to the two types of numbers—the perfect square ends with a pair of identical factors while for a number that is not a perfect square all pairs have distinct factors. As discussed later, this observation can form the basis for a generic proof that a perfect square of a natural number has an odd number of divisors.

Note that Balacheff (1988) uses the term generic example differently, that is, to describe a certain type of proof, actually the third up in the hierarchy of four types of ‘proof’ that he proposes: Naive empiricism, the crucial experiment, the generic example, and the thought experiment. It appears that Balacheff attributes a different or more restricted meaning than Mason and Pimm. Balacheff's notion of generic example falls more readily to the description of a *generic proof* (see later).

As mentioned above, the chapters in this section extend the more commonly used notion of a mathematical example to examples of meta-concepts, as mathematical arguments and proofs. It seems that Tsujiyama and Yui, Kempen, and Reid and Vallejo Vargas, are treating the examples that they present as generic meta-examples. They are using them to convey their general viewpoints.

Who Is the Source of the Example?

Zaslavsky (2014, 2017) notes that the source of an example has bearing on how a student may interpret it and what use s/he may make of it. There are differences in students' behaviour and in what they may gain from using examples, depending on whether an example is spontaneously generated, explicitly evoked, or provided as part of the task. *Spontaneous* construction of an example may indicate an inner need to communicate an idea or to make sense of the situation at hand; An *evoked* construction of an example refers to being deliberately triggered to use an example in a certain situation; and a provided example calls for *responsive* consideration in the relevant context. The findings of Aricha-Metzer and Zaslavsky (2017) from proof eliciting task-based interviews with students from a range of age levels (middle and high school as well as undergraduates), indicate that when the interviewer was the source of an example, more students were able to use the example productively for proving (compared to cases where the student was the source of the example). According to Mason (2017) these distinctions provide further insight into *scaffolding* and *fading*, and are closely related to what Love and Mason (1992) refer to as the 'gradual internalisation of an action'.

Related to the dimension of the source of an example, the four chapters in this section address different aspects. Buchbinder reminds us that there are types of tasks that lend themselves to spontaneous use of examples, as the task type used in her study: "Is this a coincidence?" (for additional illustration and discussion of spontaneous example-use elicited by a similar task, based on the mediant property, see Zaslavsky 2010). This is an example of a task that created uncertainty regarding the validity of a conjecture (Zaslavsky 2005), thus, as seen in the findings of Buchbinder, motivated students to systematically explore their initiated examples in order to gain confidence with respect to its validity.

Tsujiyama and Yui (this volume) present a provided example to their students—an example of a hypothetical student's unsuccessful argument for constructing a proof. This example aimed at raising students' reflection on their process of constructing a proof and encouraging them to analyse why this argument was unsuccessful for proving the given claim, and by this—highlighting the nature of the successful arguments. As mentioned above, this can be seen as a provided meta-example. Note that when an example is provided by a teacher or researcher, the timing may make a difference. For example, Tsujiyama and Yui decided to provide the example after the students completed the phase of constructing a proof. Similarly, Aricha-Metzer and Zaslavsky provided an example that they thought would be helpful for proving a conjecture only after the students exhausted their (unsuccessful) attempts, with or without using their own examples.

Reid and Vallejo Vargas as well as Kempen (this volume) relate to the issue of the source of an example, particularly to provided examples, by introducing the distinction between an *author* and a *reader* of an example, possibly inspired by Balacheff's (1988) term of *producer*. The *author* is the one who provides the example, with a certain intention and interpretation in mind, while the *reader* interprets it through his or her own lens. I find their distinction extremely helpful in discussing sense making and subjective aspects of example use and interpretation. This distinction is closely related to the psychological element that Reid and Vallejo Vargas discuss and to the phenomenon mentioned by Mason and Pimm (1984) of the possible "mis-match" between a teacher's intention and what students attend to in an example. I would add to the *author* and *reader* also the role of a *critic*, who can see how both the author and the reader approach an example, and may offer additional lens through which to examine the example.

The Many Facets of Example-Based Reasoning

Recently, there has been a growing body of research that focuses on example-based reasoning in mathematics, or more specifically, on the role of examples in proving-related activities (e.g., Iannone et al. 2011; Knuth et al. 2017; Sandefur et al. 2013). Example-based reasoning refers broadly to reasoning with examples, whether the examples are treated generically or not, as examples 'can provide a reasoner with a great deal of leverage' (Rissland Michener 1991). Related to example-based reasoning in the context of conjecturing and proving, Aricha-Metzer and Zaslavsky (2017) distinguish between empirical example-use and generic example-use. By *empirical example-use* they refer to the use of specific examples to make sense of, check out, or verify conjectures, focusing on the specifics of an example without looking at the specifics in a general way, while a *generic example-use* implies seeing through the specifics the general case, and includes generic proving, which is considered a particularly kind of example-based reasoning that many scholars favour (e.g., Leron and Zaslavsky 2013; Rowland 1998, 2001; Yopp and Ely 2016).

The chapters in this section of the book focus on example-based reasoning in its broad sense (Rissland Michener 1991). Research on students' example-use in the course of conjecturing and proving has focused until recently mainly on students' overreliance on examples as sufficient evidence for determining the validity of a mathematical proposition (e.g., Harel and Sowder 1998; Healy and Hoyles 2000; Iannone et al. 2011; Knuth et al. 2009). I find it interesting and encouraging that all four chapters in this section of the book take a different stand, as do other recent studies (e.g., Ellis et al. 2012; Knuth et al. 2012; Pedemonte and Buchbinder 2011) by examining ways to help students become aware of the potential value of example-based reasoning and ways to use examples in productive and valid ways, when constructing or making sense of a proof. In particular, Reid and Vallejo Vargas, and Kempen, offer an analysis of the interplay between examples and proofs, or more specifically, generic proofs, that illustrate ways in which examples can and should be used for proving. Even Buchbinder, whose interviewees exhibited empirical example-use, which reflected over-reliance on specific examples in drawing general conclusions, chose to take a closer look at the data, maintaining that "even though the overall reasoning was incorrect, students' arguments seem to involve quite sophisticated mathematical thinking. Hence, deeper analysis of responses in this category might shed light on the reasoning processes underlying students' thinking, inform our understanding of students' conceptions of proving and suggest potential mediating solutions." (ibid). In a way, all four chapters attempt to capitalize on students' example-based reasoning.

I turn to some questions arising in from the chapters in this section that relate to a special kind of example-based reasoning: generic proofs.

What Counts as a Generic Proof?

For Mason and Pimm (1984) "a generic proof, although given in terms of a particular number, nowhere relies on any specific properties of that number" (ibid, p. 284). While this description captures the essence of a generic proof, it is rather vague and does not provide a clear criterion by which to determine whether a certain manifestation of example-based reasoning constitutes a generic proof for the 'arguer' and/or for the 'listener'. The mere term *generic proof* implies that this is considered a certain type of proof. This may explain Reid and Vallejo Vargas' (this volume) use of the term generic argument instead of generic proof. Their term may also reflect the inclusion in the discussion of a wider range of arguments that use examples, as they ask "When is a generic argument a proof?".

While there seems to be an agreement in principle on the features of a generic proof, there do not seem to be well defined criteria for what counts as a generic proof, even within a given community with shared norms and goals. The main differences rest on the amount of detail and explanation needed to justify why the proof carried out on a (generic) example will work for any other one as well.

Reid and Vallejo (this volume) interpret Balacheff's notion of a generic example as a generic argument that "involves making explicit the reasons for the truth of an assertion by means of operations or transformations on an object that is not there in its own right, but as a characteristic representative of its class. The account involves the characteristic properties and structures of a class, while doing so in terms of the names and illustration of one of its representatives." (Balacheff 1988, p. 219). This concurs with Dreyfus et al. (2012), on which Kempen (this volume) relies, who distinguish between example-based arguments, example-based generic arguments and general arguments.

Unlike a generic example, a generic proof is always associated with a mathematical claim or conjecture, thus, it is expected that a generic proof *convinces* the target person (author, reader, or both) that the claim is true for all cases, unless it is a false claim, for which a counter-example is sufficient for disproving. Peled and Zaslavsky (1997) discuss the *explanatory* feature of a counter-example, which could be considered a generic (dis)-proof, as it not only disproves but also conveys the reason why the claim is false.

In addition to the element of *conviction*, it is expected that a generic proof sheds light on *why* the claim is true, as Balacheff asserts that the generic example is "no longer a matter of 'showing' the result is true because 'it works'; rather, it concerns establishing the necessary nature of its truth by giving reasons" (ibid, p. 218). With respect to the necessary nature of the truth of a statement, Leron and Zaslavsky (2013) point to the need to take into account a possible pitfall that characterizes generic proofs—when operating on an example, "some phenomena *just happen,* automatically", however, one needs to justify why this will always happen, in every single case. A careful choice of the example used in a generic proof could help address this point, as according to Leron and Zaslavsky an example that is "complex enough" may ensure that all the main ideas of the target proof will naturally surface in the context of the example. The strength of a generic proof is that it reduces the level of abstraction and suspends or even eliminates the need to deal with formalism and symbolism (Dreyfus et al. 2012; Leron and Zaslavsky 2013; Rowland 1998, 2001), and by this may facilitate the transition from inductive informal to more deductive formal reasoning (Knuth et al. 2017; Stylianides and Stylianides 2009). This includes providing a sound *explanation* why the phenomena that is observed for a certain example would work for any other one as well.

Reid and Vallejo Vargas as well as Kempen (this volume) address this issue by the requirement to be explicit about the reasons that necessitate the truth of the statement. Note, that "it is crucial to consider who is reading the proof; it is easy to imagine a proof that is explanatory to one student but not to another and a good teacher cannot overlook this difference." (Weber 2010, p. 34). Kempen's findings support Weber's assertion, as the conceptions the participants in his study manifested varied with respect to the degree of conviction and explanatory power they attributed to different proofs, and moreover, the majority did not perceive these features the same as the researcher did.

It appears that not only is examplehood in the eyes of the beholder (Zaslavsky 2014), but also "generic proof" is, as illustrated next.

How Do Students Interpret and Use Generic Proofs?

As discussed above, the notion of a generic proof has been addressed by several researchers, however, research on the actual affordances of using generic proving, as a pedagogical tool for facilitating viable arguments leading to or supporting proof, are scarce. I see two main questions that would be particularly interesting to pursue: (1) How do students generate and use generic arguments or even proofs (either by their own initiative or when evoked to do so)? (2) How helpful can generic proofs that are provided by a teacher or researcher be for students, for learning to prove?

Reid and Vallejo Vargas (this volume) address the first question, by bringing an example of a student's "generic proof using several examples to show how the structure applies to other cases" of the statement: "*The square of an even natural number is always divisible by four*".

The reason the authors consider this a generic proof, is because the square of an even number is represented by a square shaped dot pattern and is accompanied by a detailed explanation, that starts with a dot pattern of the even number is 6 and its square is 36. She divides the square of $6 \times 6 = 36$ dots into smaller squares of $2 \times 2 = 4$, thus, concludes that the number ($6^2 = 36$) is divisible by 4. Until this point, she has only established that this method of dividing the large square into $2 \times 2 = 4$ squares works for the case of $n = 6$, that is, for the square $6 \times 6 = 36$. What is missing is an argument that explains why this method would work for *any* even natural number. The student addresses this by writing: "This is true for all even n, because as the side length increases by 2 dots, so new squares encompassing 4 dots are added. As n increases by 2 (next even n), $n-1$ squares encompassing 2×2 dots are added." (ibid). It appears that the student detected a pattern by observing the examples she tried (for $n = 2, 4, 6, 8$), and noticed the structure in these cases. Yet, once she tried to reason with a general symbol n, she used the same n to denote an even number as well as the next even number (instead of $n + 2$), and formulated a formal argument. It could be argued that there is not sufficient evidence that the student understands *why* this structure would work for any even number, which is often the case when judging only by written work. This example indicates the challenge and potential strength of students' generic proving. Although we do not have sufficient background information to draw conclusions, it appears that the student in Reid and Vallejo Vargas' study used the examples to help her make sense of the statement and understand why it will always work, and at the same time to communicate her reasoning.

Within the framework of a large study on the use of examples in learning to prove (Aricha-Metzer and Zaslavsky 2017; Knuth et al. 2017; Zaslavsky et al., forthcoming), we addressed both of the above questions. In individual task-based interviews, the participants were given, amongst other conjectures, the following one, used also in Kempen's study: "*If you add any number of consecutive numbers together, the sum will be a multiple of however many numbers you added up*".

They were asked if they thought the conjecture was true for any 5 consecutive numbers, and why they thought what they did. This is a rich task and lends itself for generic proving as well as disproving (depending on the parity of the number of consecutive integers that are added).

Students who did not find a way to justify (or falsify) this conjecture, were offered the following prompt, that was meant as the basis for a generic proof:

> Another student had an idea of how to explain this conjecture. For the *five* consecutive numbers *5, 6, 7, 8, 9*, she decided to write the sum as:
>
> $(7-2)+(7-1)+7+(7+1)+(7+2)$,
>
> and writing it that way helped her explain why the sum must be a multiple of 5.
>
> How do you think that helped her see why the rule is true for any five consecutive numbers? (Aricha-Metzer and Zaslavsky 2017).

Interestingly yet not surprisingly, there were students who immediately were able to see through this example the general case and provide a sound explanation why this will work for any five consecutive numbers, as one of the interviewees Roger said: "Okay, so I can see now—this is pretty good proof for why it's, uh, for why it has to be a multiple of 5 or just a multiple of an odd number in general…." (ibid). Roger goes on and explains in detail why the conjecture works not only for five consecutive numbers, but also for any sum of an odd number of consecutive numbers. Moreover, this example also helped him see why the conjecture does not hold for an even number of addends, as he noticed the role of the middle term and the symmetry around it.

On the other hand, Daniel, who was offered the same kind of generic example for the sum of 2 + 3 + 4 + 5 + 6, still did not see the general idea of the proof, as seen from his response: "…That helped her understand it better? [This] is just another way of rewriting 2 + 3 + 4 + 5 + 6. I don't understand how it helped her explain it like this…".

Along the same lines, findings reported by Kempen (this volume) show that a vast majority of the participants in his study did not attribute the same explanatory power as he did to the proofs that he considered generic.

How Generic Is a Generic Proof?

A generic proof may convey only some but not all the ideas of a proof, or may not be equally transparent to all the main ideas of the proof.

Kempen (this volume) brings what he considers a "generic proof in the context of figurative numbers" of the statement: "*The sum of five consecutive numbers is always divisible by five*" (Fig. 20.1).

For Kempen as well as for Reid and Vallejo Vargas, a generic proof should include explicit verbal or other means of justification why the method/result will always hold for any case. Along these lines, Kempen (this volume) adds a verbal explanation to Fig. 20.1.

One question that comes to mind is how insightful this generic proof is for the more general case. I would say that Kempen's generic proof of the sum of consecutive numbers conveys the reason why the statement works for any five

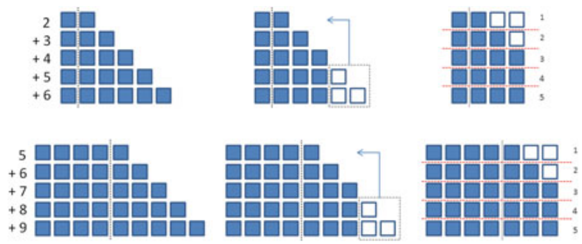


Fig. 20.1 "The sum of five consecutive numbers represented by figurate numbers" (Kempen, this volume, p. 230)

consecutive numbers, but does not convey what the role of the oddness of five plays, or what Roger was able to see through the generic example that was offered to him (see above)—that is, why the conjecture works for the sum of *any* odd (but not even) number of consecutive numbers. This implies, that from the perspective of the author of the generic proof, in particular if he or she are the teacher, there are additional considerations regarding the choice and design of a generic proof.

The different levels of genericity of a generic argument or proof is also addressed by Reid and Vallejo Vargas (this volume). For example, in the third generic argument that they present for the claim: "*The sum of the first n positive integers is* $\frac{n(n+1)}{2}$", they separate between the case when n is odd and when it is even. For an odd n they write:

> First, consider an odd n, for example 7. Then the sum is $1+2+3+4+5+6+7$. You can rearrange this is to 3 pairs: $1+7, 2+6, 3+5$, all adding up to 8, with the 4 in the middle left out. So, the sum is $3 \times 8 + 4$, or in general $\left(\frac{n-1}{2}\right)(n+1) + \left(\frac{n+1}{2}\right)$, which simplifies to $\left(\frac{n(n+1)}{2}\right)$. (Reid and Vallejo Vargas, this volume, p. 242)

The authors recognize that this argument makes it more difficult to understand why in the general formula we divide by 2 or multiply by $(n+1)$, compared, for example, to the first argument they bring, which is more transparent to the general formula.

Similarly, the 'generic proof with numbers' that Kempen presents to the statement that: "*The sum of an odd natural number and its double is always odd*" is also limited in its genericity:

> $1 + 2{\cdot}1 = 3{\cdot}1 = 3$, $5 + 2{\cdot}5 = 3{\cdot}5 = 15$, $13 + 2{\cdot}13 = 3{\cdot}13 = 39$
>
> The sum of an odd natural number and its double equals three times the initial number. Since the initial number is an odd number, one obtains the product of two odd numbers. Since the product of any two odd numbers is always odd, the result will always be an odd number. (Kempen, this volume, p. 230, original in italics)

The given examples do not convey *why* the product is odd, only why the sum is three times the initial number.

When Is There a Need for a Generic Proof?

The necessity principle for learning and in particular with respect to learning to prove (e.g., Harel 2013; Zaslavsky et al. 2012) is an important one that should guide many decisions and choices of teachers. In reading the chapters, I was contemplating to what extent would it make sense and would it be necessary to use a generic (figurative or other) proof, rather than a general/deductive proof. Clearly, the necessity varies from one audience to another. As a rule, I would expect a generic proof to be particular helpful either when a more general/deductive proof is not accessible (for example, because it is too complicated and abstract), and/or when an appropriate notation is either cumbersome or non-existent. By this criteria, Reid and Vallejo Vargas' generic arguments for the sum of the first n natural numbers seem to stem from a genuine necessity, while Kempen's generic proof of the sum of an odd natural number and its double may not be equally necessary, as at a rather early stage, students are able to understand that if you take a number and add to it its double, you get three times the number. Moreover, this part would work for any number, not just positive integers. Basically, there are two main ideas in the general proof—one is based on the fact that this sum is three times the number, and the second is based on the fact that the product of two odd integers is an odd integer. Kempen's proof does not offer any insight to the latter.

To get a sense of a strong need for a generic proof, consider the following statement: "*A natural number has an odd number of factors* ***if and only if*** *it is a perfect square.*"

Constructing a formal proof of this statement can get rather messy, partly because the issue of notation for the general case is not at all trivial. I have tried this out with several groups of undergraduate math majors and secondary math teachers, and noted that many of them reached an impasse. Even a number of mathematicians who were given this statement, started out and sufficed with generic examples. Apparently, this statement lends itself well for a generic proof.

I used the number 36 earlier, as a generic example for a *procedure* to find all the factors of a number, by listing systematically all its factorizations as a product of two factors. For many students at all levels, that mere presentation led to an "Aha" moment with respect to the above statement, where they felt that they instantly gained insight into the statement and to why it is true for any natural number, that is, what about a perfect square guarantees that it has an odd number of factors, but also why any natural number that is not a perfect square must have an even number of factors. In Reid and Vallejo Vargas' term (this volume), both the 'author' and many of the 'readers' were able to see through the generic example the main idea of the proof and to articulate it.

Note that the conjecture posed to the participants in Buchbinder's study, does not seem to lend itself to generic proving, at least not in its mode of representation. This may explain why the students did not use examples generically, rather systematically. There is a different way to present the mediant property in a way that lends itself to (graphical) generic proving (Zaslavsky 2010). It would be interesting to see if and how students' responses would change accordingly.

On Rigor and Evidence

I conclude with the overarching questions of the place of rigor and evidence in example-based reasoning.

Reid and Vallejo Vargas refer to what they consider a debate regarding the status of a generic proof, that is, whether it is acceptable as a mathematical proof. They raise the issue of "fussiness" (which basically has to do with the degree of rigor that is expected) discussed by Leron and Zaslavsky (2013), as a major source of this supposed debate. This view ignores the context of the work of Leron and Zaslavsky, and the wealth of aspects and considerations discussed in their paper. Leron and Zaslavsky's views complement rather than contradict other views in the field.

As in the status of general proofs, the norms for the degree and kind of rigor that is required for acceptance of a (generic) proof depend on several elements, such as the community—whether it is a community of professional mathematicians, a classroom of a certain grade-level, or other. Even within the mathematics community there is no consensus regarding the level of rigor that is required. This relates to the social factor that Reid and Vallejo Vargas discuss. Stylianides (2007) conceptualized the meaning of proof in school mathematics, maintaining the essence of the meaning of proof in the mathematics community, with modifications. Clearly the warrants that are expected in school are different than those at the university.

The degree of rigor that is expected in a proof depends not only on the community but also on the level of sophistication of the proof. Some generic proofs are more transparent and convincing than others. For example, it can be argued that in Leron and Zaslavsky's (2013) Case Study 1, (i.e., the generic proof that a perfect square of a natural number has an odd number of distinct factors) it is easy to explain why the method used for the generic example (36) will work for any other perfect square (as explained earlier in this paper), thus, there is no need for further rigor. However, in Case Study 2 (ibid), (i.e., the generic proof that every permutation has a unique decomposition as a product of disjoint cycles) it is not obvious why the same phenomena will occur for a different example, thus, there is a need for more rigor in order to prove the statement. Yet, even mathematicians are urged to keep in mind that the main ideas of the proof need to be revealed, as "it will not do to bury the idea under the formalism" (Mac Lane 1986, p. 378).

Similar to the subjective aspect of genericity that was discussed earlier, there is a subjective aspect to the acceptance of a generic proof as a proof, even within the same community of observers (Pauletti and Zaslavsky, in preparation). This subjectivity is also connected to the diverse views Kempen (this volume) found regarding what he terms verification and Weber and Mejía-Ramos (2015) term absolute conviction. It appears that for his participants a formal proof was more convincing.

As researchers and teachers, we are constantly seeking strong evidence for our interpretations of students' understanding, and often feel that the evidence is not sufficiently compelling and that there is some degree of speculation in our interpretations. In order to accept a student's example-based argument as a proof—Reid and Vallejo Vargas require certain kind of evidence indicating that the student is aware of the generality of the argument beyond the example that is used, and that he or she can articulate why the claim at hand must hold in the same way and for the same reasons for any other case/example, based on the norms and conventions of the classroom community. Argument 1[b] (ibid) is a good example of the kind of evidence needed. The latter can be seen as the "fussiness" dictated by the community. The question is—who is the author and who is the reader, and is this evidence needed for both?

With respect to the issue of evidence, in her chapter, Buchbinder interprets the participants' systematic exploration of examples as an over-generalization of a method they had learned in another context. However, there is no sound evidence that the students actually learned this method of solving inequalities by dividing the domain to sub-domains and checking one number from each sub-domain, and even if they did—they may not have made this connection. Kempen, draws conclusions from the responses to a written questionnaire regarding the conceptions held by the participants in his study about the levels of the explanatory power, subjective conviction, and absolute conviction (what he terms verification) of each of four types of proofs. It is not clear what meaning the participants attributed to these three constructs, and what features of the different proofs they attended to. Tsujiyama and Yui bring evidence that Mizu reflected on her proving plan and gained insight based on the imaginary student's (Mikio) unsuccessful arguments. Would she have gained the same insight if she had not reached a similar impasse on her own?

These questions remain open and call for further investigations and refinements of our criteria for both rigor and evidence.

References

Alcock, L., & Weber, K. (2010). Referential and syntactic approaches to proving: Case studies from a transition-to-proof course. In F. Hitt, D. Holton, & P. Thompson (Eds.), *Research in collegiate mathematics education VII* (pp. 93–114). Washington: AMS.

Aricha-Metzer, I., & Zaslavsky, O. (2017). The nature of productive and non-productive example-use in proving. *Journal of Mathematical Behavior*. http://dx.doi.org/10.1016/j.jmathb.2017.09.002.

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 and Stoughton.

Dreyfus, T., Nardi, E., & Leikin, R. (2012). Forms of proof and proving in the classroom. In G. Hanna & M. de Villiers (Eds.), *Proof and proving in mathematics education* (pp. 191–214). New York: Springer.

Ellis, A. B., Lockwood, E., Williams, C., Dogan, M. F., & Knuth, E. (2012). Middle school students' example use in conjecture exploration and justification. In L. R. Van Zoest, J. J. Lo, & J. L. Kratky (Eds.), *Proceedings of the 34th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 135–142). Kalamazoo, MI: Western Michigan University.

Goldenberg, P., & Mason, J. (2008). Shedding light on and with example spaces. *Educational Studies in Mathematics, 69*(2), 183–194.

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

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Research in collegiate mathematics education III* (pp. 234–283). Washington, DC: Mathematical Association of America.

Harel, G., & Tall, D. (1991). The general, the abstract, and the generic in advanced mathematics. *For the Learning of Mathematics, 11*(1), 38–42.

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education, 31*(4), 396–428.

Hershkowitz, R. (1990). Psychological aspects of learning geometry. In P. Nesher & J. Kilpatrick (Eds.), *Mathematics and cognition: A research synthesis by the International Group for the Psychology of Mathematics Education* (pp. 70–95). NY: Cambridge University Press.

Hershkowitz, R., & Vinner, S. (1983). The role of critical and non-critical attributes in the concept image of geometrical concepts. In R. Hershkowitz (Ed.), *Proceedings of the 7th Conference of the International Group for the Psychology of Mathematics Education* (pp. 223–228). Weizmann Institute, Israel: PME.

Iannone, P., Inglis, M., Mejía-Ramos, J. P., Simpson, A., & Weber, K. (2011). Does generating examples aid proof production? *Educational Studies in Mathematics, 77,* 1–14.

Knuth, E., Choppin, J., & Bieda, K. (2009). Middle school students' production of mathematical justifications. In D. Stylianou, M. Blanton, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K–16 perspective* (pp. 153–170). New York, NY: Routledge.

Knuth, E., Kalish, C., Ellis, A., Williams, C., & Felton, M. (2012). Adolescent reasoning in mathematical and non-mathematical domains: Exploring the paradox. In V. Reyna, S. Chapman, M. Dougherty, & J. Confrey (Eds.), *The adolescent brain: Learning, reasoning, and decision making* (pp. 183–209). Washington, DC: American Psychological Association.

Knuth, E., Zaslavsky, O., & Ellis, A. (2017). The role and use of examples in learning to prove. *Journal of Mathematical Behavior*. http://dx.doi.org/10.1016/j.jmathb.2017.06.002.

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

Love, E., & Mason, J. (1992). *Teaching mathematics: Action and awareness*. Milton Keynes: Open University.

Mac Lane, S. (1986). *Mathematics form and function*. New York, NY: Springer.

Mason, J. (2017). Relationships between proof and examples. *Journal of Mathematical Behavior*. http://dx.doi.org/10.1016/j.jmathb.2017.07.005.

Mason, J., & Pimm, D. (1984). Generic examples: Seeing the general in the particular. *Educational Studies in Mathematics, 15*(3), 277–290.

Mills, M. (2014). A framework for example usage in proof presentations. *Journal of Mathematical Behavior, 33,* 106–118.

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

Pauletti, K., & Zaslavsky, O. (in preparation). Mathematicians' views on the acceptance of generic arguments of students.

Pedemonte, B., & Buchbinder, O. (2011). Examining the role of examples in proving processes through a cognitive lens: The case of triangular numbers. *ZDM: The International Journal on Mathematics Education, 43*(2), 257–267.

Peled, I., & Zaslavsky, O. (1997). Counter-examples that (only) prove and counter-examples that (also) explain. *FOCUS on Learning Problems in Mathematics, 19*(3), 49–61.

Reiss, K., Heinze, A., Renkl, A., & Groβ, C. (2008). Reasoning and proof in geometry: Effects of a learning environment based on heuristic worked-out examples. *ZDM The International Journal on Mathematics Education, 40*(3), 455–467.
Rissland Michener, E. L. (1978). Understanding understanding mathematics. *Cognitive Science, 2* (4), 361–383.
Rissland Michener, E. L. (1991). Example-based reasoning. In J. F. Voss, D. N. Parkins, & J. W. Segal (Eds.), *Informal reasoning in education* (pp. 187–208). Hillsdale, NJ: Lawrence Erlbaum Associates.
Rowland, T. (1998). Conviction, explanation, and generic examples. In A. Olivier & K. Newstead (Eds.), *Proceedings of the 22nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 65–72). Stellenbosch, South Africa: University of Stellenbosch.
Rowland, T. (2001). Generic proofs in number theory. In S. Campbell & R. Zazkis (Eds.), *Learning and teaching number theory: Research in cognition and instruction* (pp. 157–184). Westport, CT: Ablex Publishing.
Sandefur, J., Mason, J., Stylianides, G. J., & Watson, A. (2013). Generating and using examples in the proving process. *Educational Studies in Mathematics, 83*(3), 323–340.
Skemp, R. R. (1987). *The psychology of learning mathematics: Expanded American edition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Stylianides, A. J. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education, 38*(3), 289–321.
Stylianides, G. J., & Stylianides, A. J. (2009). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education, 40*(3), 314–352.
Vinner, S. (1983). Concept definition, concept image and the notion of function. *International Journal of Mathematical Education in Science and Technology, 14,* 293–305.
Vinner, S., & Dreyfus, T. (1989). Images and definitions for the concept of function. *Journal for Research in Mathematics Education, 20*(4), 356–366.
Watson, A., & Mason, J. (2005). *Mathematics as a constructive activity: Learners generating examples*. Mahwah: Erlbaum.
Weber, K. (2010). Proofs that develop insight. *For the Learning of Mathematics, 30*(1), 32–36.
Weber, K., & Mejía-Ramos, J. P. (2015). On relative and absolute conviction in mathematics. *For the Learning of Mathematics, 35*(2), 15–21.
Yopp, D. A., & Ely, R. (2016). When does an argument use a generic example? *Educational Studies in Mathematics, 91,* 37–53.
Zaslavsky, O. (2005). Seizing the opportunity to create uncertainty in learning mathematics. *Educational Studies in Mathematics, 60,* 297–321.
Zaslavsky, O. (2010). The explanatory power of examples in mathematics: Challenges for teaching. In M. K. Stein & L. Kucan (Eds.), *Instructional explanations in the disciplines* (pp. 107–128). New York: Springer.
Zaslavsky, O. (2014). Thinking with and through examples. In P. Liljedahl, C. Nicol, S. Oesterle, & D. Allan (Eds.), *Proceedings of the Joint Meeting of PME 38 and PME-NA 36,* (Vol. 1, pp. 21–34). Vancouver, Canada: PME.
Zaslavsky, O. (2017). There is more to examples than meets the eye: Thinking with and through examples in different settings. *Journal of Mathematical Behavior*. http://dx.doi.org/10.1016/j.jmathb.2017.10.001.
Zaslavsky, O., Knuth, E., & Ellis, A. (Eds.) (forthcoming). The roles of examples in proving and learning to prove: Special Issue. *Journal of Mathematical Behavior.*
Zaslavsky, O., Nickerson, S., Stylianides, A., Kidron, I., & Winicki-Landman, G. (2012). The need for proof and proving: Mathematical and pedagogical perspectives. In G. Hanna & M. de Villiers (Eds.), *Proof and proving in mathematics education* (pp. 215–229). New York: Springer.
Zaslavsky, O., & Shir, K. (2005). Students' conceptions of a mathematical definition. *Journal for Research in Mathematics Education (JRME), 36*(4), 317–346.

Index

A. J. Stylianides and G. Harel (eds.), *Advances in Mathematics Education Research on Proof and Proving*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70996-3