View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Loughborough University Institutional Repository

Loughborough University

This item was submitted to Loughborough's Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

creative commons

COMMONS DEED

Attribution-NonCommercial-NoDerivs 2.5

You are free:

- to copy, distribute, display, and perform the work

Under the following conditions:

Attribution. You must attribute the work in the manner specified by the author or licensor.

Noncommercial. You may not use this work for commercial purposes.

No Derivative Works. You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

Your fair use and other rights are in no way affected by the above.

This is a human-readable summary of the Legal Code (the full license).

Disclaimer

For the full text of this licence, please go to:
http://creativecommons.org/licenses/by-nc-nd/2.5/

How mathematicians obtain conviction:

Implications for mathematics instruction and research on epistemic cognition

Keith Weber[1], Matthew Inglis[2] & Juan Pablo Mejia-Ramos[1]

[1]Rutgers University [2]Loughborough University

Abstract. The received view of mathematical practice is that mathematicians gain certainty in mathematical assertions by deductive evidence rather than empirical or authoritarian evidence. This assumption has influenced mathematics instruction where students are expected to justify assertions with deductive arguments rather than by checking the assertion with specific examples or appealing to authorities. In this paper, we argue that the received view about mathematical practice is too simplistic; some mathematicians sometimes gain high levels of conviction with empirical or authoritarian evidence and sometimes do not gain full conviction from the proofs that they read. We discuss what implications this might have, both for for mathematics instruction and theories of epistemic cognition

Key words: Conviction; Epistemic cognition; Justification; Mathematics; Mathematics education; Proof

One of the central problems in contemporary mathematics education concerns how students should convince themselves and others that mathematical assertions are true (Harel & Sowder, 1998, 2007). For instance, the Common Core Standards list a number of skills related to justification[1] that students should develop in elementary and high school school mathematics, including “comparing the effectiveness of two plausible arguments,” “distinguishing correct reasoning from that which is flawed,” and “deciding whether [the arguments] of others make sense” (Common Core Content Standards Initiative, 2012). It is widely assumed by mathematics educators that knowledge should be constructed and justified in mathematics classrooms in a manner analogous to how knowledge is constructed and justified in the mathematical community (Ball & Bass, 2000; Schoenfeld, 1992) and that students should gain conviction and persuade others in the same ways that mathematicians do (Harel & Sowder, 1998, 2007). Of course, there will necessarily be important differences between a classroom community and a mathematical community. For instance, the concepts that mathematicians study are often more abstract and further removed from direct perception; mathematicians also have access to conceptual resources, such as powerful representation systems and methods of inference, with which students are unfamiliar. Nonetheless, mathematics educators argue that students’ standards of conviction and justification can still be qualitatively similar to those of mathematicians in that they can use the same types of evidence to justify their mathematical claims. The common assumption is that like mathematicians, students

[1] Throughout the paper, we refer to justification as any argument that aims to establish the veracity of a claim.

should be led to use deductive reasoning[2], rather than empirical reasoning or appeals to authority, to justify their mathematical knowledge (Harel & Sowder, 1998, 2007; Recio & Godino, 2001; Stylianides, 2007). The desire is for students to engage in similar epistemic actions as mathematicians when it comes to justifying mathematical assertions.

Stylianou (2004) argued that "if indeed we choose to focus [instruction] on professional mathematical practice, then it is crucial to have a thorough understanding of this practice" (p. 304). We concur with this statement. If a central goal of mathematical instruction is for students to convince themselves and others of the truth of an assertion using the same types of evidence that mathematicians do (Harel & Sowder, 2007), then it is essential to have an accurate understanding of how mathematicians convince themselves that mathematical assertions are true. Put another way, if we want mathematicians' epistemic cognition to be an explicit goal of mathematics instruction, research is needed on what expert cognition in mathematics is. We believe that recent arguments in the philosophy of mathematics, as well as the results of many empirical studies that we have conducted, imply the position that mathematicians gain conviction solely by logical deduction is too simplistic. In this paper, we will argue that a more nuanced view of mathematical practice is necessary and that this has important consequences for mathematics instruction. Specifically, we will advance the following claims:

[2] Throughout the paper, when we refer to deductive reasoning, we are referring to arguments where each new assertion is a necessary logical consequence of previous assertions, not where each statement is derived by the mechanical application of an explicit logical rule. Indeed, few proofs in mathematics are written in this way. We discuss this point in more detail later in this paper.

(1) The position that mathematicians never gain conviction in mathematical assertions from empirical evidence or authoritarian sources is overstated.

(2) The sources of evidence that mathematicians use are dependent upon their epistemic aims, a point that has until now received only limited consideration in the mathematics education literature. Mathematicians do not believe that an argument based on empirical evidence constitutes a mathematical proof [3], but this does not imply that they never find empirical evidence to be persuasive.

(3) Whether empirical or authoritarian evidence is persuasive to a mathematician depends, in part, on the quality of that evidence. Simply because mathematicians reject some types of empirical or authoritarian evidence as unpersuasive does not imply that all other types of empirical or authoritarian evidence are never persuasive.

(4) Whether an argument is acceptable to a mathematician depends, in part, upon the individual mathematician. There is substantial variance in mathematicians' evaluations of mathematical arguments. This is true for mathematicians' evaluations of both empirical and deductive arguments.

(5) The instructional goal of having students be persuaded by deductive arguments, but not empirical evidence or appeals to authority, has merit. However, the simplicity of this viewpoint leads to some instructional goals that we find to be both infeasible and

[3] There is no consensus on the meaning of proof, either among philosophers (e.g., Rav, 2007) or mathematics educators (e.g., CadwalladerOlsker, 2011), although we believe the majority of mathematics educators believe any characterization of proof should be based on the proofs that mathematicians actually produce (CadwalladerOlsker, 2011). For the purposes of this paper, we define a proof as a deductive argument that claims to show a conclusion is a logically necessary consequence from agreed upon assumptions. One contribution of this paper is exploring what types of arguments that mathematicians consider to be deductive and to be proofs.

undesirable. We suggest more nuanced instructional recommendations as the result of the arguments we present in this paper.

(6) Most models of epistemic cognition are based on continuums contrasting undesirable epistemic beliefs with more sophisticated beliefs. Correspondingly, researchers contrast unproductive epistemic actions with more productive ones. The goal of instruction is then to have students reject poor epistemic beliefs and cease using unproductive epistemic actions and adopt more sophisticated beliefs and actions. We will argue that some purportedly undesirable epistemic beliefs and actions actually are appropriate in some circumstances for certain epistemic aims. We suggest a somewhat different goal of instruction for epistemic cognition: Rather than replace purportedly poor beliefs with better ones, it would be better to (a) expand the repertoire of epistemic actions available to a student and (b) develop in the student an appreciation both the limitations of different types of epistemic actions as well as situations in which they are useful and appropriate to invoke.

(6) Recently, there has been interest in how epistemic cognition differs across domains (e.g., Greene, Torney-Puta, & Azevedo, 2010) with a particular interest in epistemic cognition in mathematics. Yet surprisingly, this research has not been informed by the careful study of mathematicians. The results of this paper will demonstrate that several received views about mathematicians' epistemic beliefs are inaccurate. Consequently, we call for more research in expert practice to inform domain-specific models of epistemic cognition

How expert practice informs instruction

This paper is based on the premise that knowledge of how mathematicians obtain conviction can and should inform the goals and methods of mathematics instruction. Before doing so, we acknowledge that the link between expert practice and instructional design is not straightforward. We believe there are four reasons why research on expert practice is pertinent to instructional design.

(i) As noted above, many mathematics educators treat it as axiomatic that having students adopt mathematicians' beliefs and practices with respect to proof should be a primary goal of mathematics instruction. If the goal of instruction is to have students engage in experts' practices, then it is necessary to have an accurate understanding of what those practices are.

(ii) Some mathematics educators challenge the notion that mathematics classrooms should mirror mathematical practice. Staples, Bartlo, and Thanheiser (2012) argued that because the needs of a classroom community differ from those of a mathematical community, one would expect the role of justification and proof to differ within those communities as well. We agree that it might be appropriate for teachers and students to treat justification and proof differently than mathematicians if there are good pedagogical reasons for doing so. However, we also agree with Herbst and Balacheff (2009) that proof in the classroom should be "*accountable to*, if not compatible with, how it [proof] exists in the discipline" (p. 43, italics are our emphasis). To summarize our viewpoint, if justification and proving practices in a classroom differ from those of the mathematical community, the teacher should both be aware of this and have a reason for allowing or encouraging this inconsistency. Recognizing such inconsistencies requires having an accurate understanding of mathematical practice. Even if it is not problematic for students to be ignorant of mathematical practice, it does not follow that it is acceptable for teachers, researchers, or curriculum designers to be similarly unaware.

(iii) The link between expert practice and instruction is not straightforward. Indeed, there are cases of instruction based on an accurate understanding of expert practice that resulted in little or even negative gains in achievement (e.g., see Schoenfeld, 1985, chapter 3). All an understanding of expert practice can provide is a useful starting point for instructional design, where this instruction should be assessed and refined based on empirical studies. Nonetheless, it is worth emphasizing that using expert practice to inform pedagogy has led to the design of effective instruction in mathematics (e.g., Schoenfeld, 1985) and science (e.g., Chi & Van Lehn, 2012). Wineberg's (1991) study of how historians read text is particularly illustrative of how studying expert practice can inform the goals and design of instruction. In his study, Wineberg found that students treated historical text as an unproblematic account of truth while historians viewed the accounts in all texts as influenced by the perspectives of their authors. Thus, unlike students, historians considered the source of a text in interpreting its meaning and accuracy. After Wineberg's influential article was published, history educators recognized that having students consider the source of the arguments was important for interpreting texts and worked toward developing instruction that improved students' ability to do so (e.g., Britt & Aglinskas, 2002).

(iv) Most importantly, some instruction based on *purported but inaccurate conceptions of expert practice* has been harmful to students. The New Math movement in the 1970s was founded on the inaccurate conception that mathematicians' reasoning was based predominantly on set-theoretic axioms and formal notation, despite the fact that mathematicians at the time "decried [the] empty abstraction and rigid formalism" in this instruction (e.g., Wu, 1997, p. 246). New Math is now regarded as a failure because it did not focus on the understandings that mathematicians attached to their formalism. Until recently, science classrooms preached that investigations should follow the "scientific method," where the investigator proposes a hypothesis, conducts a study in which that hypothesis is tested, and then either rejects the hypothesis or strengthens his or her belief in the hypothesis based on the results of this study. As Hodson (1996) argued, this instructional paradigm is based on an inaccurate understanding of scientific practice and has been quite harmful to students, giving them a distorted view of science and leaving them unable to interpret the scientific evidence and claims that they read about in their day-to-day lives (see also Osborne et al, 2003). In this paper, we will argue that some instructional recommendations with respect to proof are based on an inaccurate view of mathematical practice and might also have undesirable consequences.

Justification and proof in mathematics education

Justification and proof are central to the discipline of mathematics. Accordingly, numerous influential mathematics educators and organizations in mathematics education have argued that justification and proof should also play a central role in all mathematics classrooms (e.g., Dreyfus, 1990; Hanna, 1995; NCTM, 2000; Schoenfeld, 1994). Further, they believe that justification and proof should not be viewed as an isolated topic in its own right or only covered in high school geometry, but should permeate the K-12 mathematics curriculum (e.g., NCTM, 2000; Schoenfeld, 1994; Stylianides, 2007). Students should continually be involved in the practice of justifying their mathematical claims.

There is a vast literature in mathematics education on justification and proof; reviews of this literature have been offered by Harel and Sowder (2007), Mariotti (2006), Weber (2003), and Yackel and Hanna (2003). In the survey of the literature in this paper, we first discuss how many mathematics educators view students' difficulties with proof as an epistemological problem and then summarize what is known about how students consider authoritarian, empirical, and deductive evidence.

Early mathematics education research on justification and proof was primarily concerned with the extent to which students could produce proofs and the cognitive processes that students needed to construct a justification or proof. Several large-scale studies (Coe & Ruthven, 1994; Healy & Hoyles, 2000; Senk, 1989) and numerous smaller ones (e.g., Bell, 1976; Moore, 1994; Weber, 2001) demonstrated that students of all ages have serious difficulties constructing proofs. For instance, Senk (1989) asked 1520 students from a high school geometry class to prove four theorems, two of which required only a single deduction beyond the given information. She found that only 30% of the sample were able to prove three of these four theorems and 29% of the sample were unable to produce a single proof. Similar results were found for students at all levels of their education, including mathematics majors completing advanced mathematics courses (e.g., Weber, 2001; Weber & Alcock, 2004). Qualitative analysis has revealed that one reason for this difficulty is that students lack the proving strategies and heuristics to write proofs effectively (e.g., Moore, 1994; Schoenfeld, 1985; Weber, 2001). Consequently some researchers have outlined the heuristics and processes that experts use to successfully write proofs (e.g., Koedinger & Anderson, 1990; Schoenfeld, 1985; Weber, 2001), and instruction designed to explicitly teach students these heuristics and processes have substantially increased students' abilities to write proofs (e.g., Anderson et al, 1994; Schoenfeld, 1985; Weber, 2006).

Reconceptualizing research on justification and proof as the epistemic commitments of students and mathematicians

In the 1990s, research on justification and proof in mathematics education shifted away from studying the processes that would enable students to produce successful proofs and toward understanding how students sought conviction in mathematical settings. This shift was precipitated by several factors, including mathematics educators moving from information-processing to constructivist theories of mind and the corresponding shift in their research focus from problem solving to conceptual understanding and task representation. Other mathematics educators became less concerned with the skills and facts that mathematics students knew and more concerned with how they participated in authentic mathematical activity (e.g., Lampert, 1990; Schoenfeld, 1992; Sfard, 1998).

Harel and Sowder (1998) published an influential paper that crystallized a research program focusing on how students sought mathematical conviction. They defined a *proof* as an argument that an individual uses to render a mathematical conjecture into a fact; that is, a proof is used to remove all personal doubts or to persuade others that the conjecture is true. An individual's *proof scheme* represents how they produce these proofs with a specific emphasis on the types of evidence they use to become certain that mathematical assertions are true. In Harel and Sowder's (1998) framework, what constitutes a proof and a proof scheme is subjective, relativistic, and student-centered. Nonetheless, Harel (2001) emphasized that "the goal of instruction must be unambiguous—to gradually refine current students' proof schemes toward the proof scheme practiced by mathematicians today" (p. 188). Since the publication of Harel and Sowder's (1998) seminal paper, their approach has become the dominant paradigm for investigating justification and proof in mathematics education. Indeed, in their review of the literature, the authors argued that most contemporary research in this area can be viewed within their framework (Harel & Sowder, 2007).

From our perspective, Harel and Sowder's contribution reconceptualized how mathematics educators perceived students' difficulties with justification and proof; instead of being understood as a deficiency in problem-solving, mathematics educators viewed students' shortcomings as an issue of epistemic cognition. The goal of mathematics instruction is to align students' epistemic cognition in mathematics with that of mathematicians, with a specific emphasis on the types of evidence[4] used to gain certainty in mathematical claims. It is assumed that if this discrepancy is addressed, students' proofs, or the epistemic products of their reasoning, will be more similar to the proofs that mathematicians produce.

Harel and Sowder (1998) presented a comprehensive taxonomy of the many types of evidence that students might use to gain complete conviction in a mathematical assertion. In this paper, we focus on three of the proof schemes that they posited that have been especially influential in mathematics education. A student with an *authoritarian proof scheme* believes a mathematical claim is true when an authoritative source such as a teacher or textbook says this is so. A student with an *empirical inductive proof scheme* believes a claim about an infinite number of objects is true when he or she verified it for a

[4] Throughout the paper, by "evidence," we mean any data that an individual regards as supporting an assertion.

finite number of specific examples. A student with a *deductive proof scheme* believes a claim is true when he or she can see how this is a necessary logical deductive consequence from other claims that are known to be true. Within this framework, the goal of instruction is to have students reject authoritarian and empirical inductive proof schemes and adopt a deductive proof scheme because this is representative of how mathematicians think and behave (Harel & Sowder, 2007).

To illustrate the types of evidence sought by a student holding each proof scheme, consider the following justifications that a student might give to justify the claim that the sum of two odd numbers is even:

Authoritarian: Our textbook tells us that this is a theorem on page 233. Since the mathematics in the textbook is correct, I believe this claim is true.

Empirical: I checked this claim with several examples. 3+5=8. 7+7 =14. 9+1 =10. 13+3 = 16. 17+11 = 28. No matter which odd numbers I tried, their sum was always even. So I believe the claim is true.

Deductive: If we have an odd number of objects, we can pair up all of these objects and have one object left over. If we have an even number of objects, we can pair up all these objects and have nothing left over. For the sum of two odd numbers, we can pair up all of the objects for the first number and have one left over, then pair up all of the objects for the second number and have one left over, then pair up the two objects that were left over. Since all the objects are now paired up, the resulting sum is even (adapted from Ball & Bass, 2000).

Note the first argument is simply an appeal to authority apart from the individual's reasoning, the second relies on the inspection of examples, and the third on logic and deduction.

Authoritarian Evidence

The role that authorities should and do play in students' construction of knowledge extends beyond justification and proof in mathematics. Indeed, this issue is considered a central one in the broader topic of epistemic cognition. Most models of epistemic cognition include a dimension measuring the extent to which students believe that knowledge resides in authorities or is constructed by the knower (e.g., Hofer & Pintrich, 1997; King & Kitchener, 1994; Muis & Franco, 2009; Perry, 1970; Schommer, 1990). In her summary of this research, Hofer (2000) claimed that, "at lower levels of most models, knowledge originates outside the self and resides in external authority, from whom it may be transmitted. The evolving conception of self as knower, with the ability to construct knowledge in interaction with others, is a development turning point of most models" (p. 381). In general, researchers studying epistemic cognition regard the belief that one should generate one's own knowledge as "sophisticated" and a reliance on authority is considered to be "naïve," in part because the former is correlated with more productive strategies for learning and more success in school (see Muis, 2004), although this viewpoint has been challenged recently (e.g., Bromme, Kienhues, & Porsch, 2009; Chinn, Buckland, & Samarapungavan, 2011).

In her review of the literature on epistemological beliefs in mathematics, Muis (2004) highlighted three main findings on beliefs with respect to authority. First, most students "believe mathematical knowledge is passively handed to them by some authority figure, typically a teacher or textbook author, and that they are incapable of learning mathematics through logic and reason" (p. 330). Second, such beliefs have a harmful effect on students' mathematical learning, leading students to focus on memorization and stifling understanding and creativity (e.g., Garafalo, 1989; Schoenfeld, 1985). Third, "it appears that environmental factors such as classroom context influence students' beliefs about mathematics" (Muis, 2004, p. 338). Harel and Rabin (2010) linked a number of common classroom practices to students' adoption of an authoritarian proof scheme, such as the teacher being the sole arbiter when there is a debate in the classroom and the teacher telling students if they are right or wrong without providing mathematical justification to support the evaluation.

In recent years, a number of researchers have sought to create instructional environments that encouraged students to build and justify their own knowledge rather than receive the knowledge from a teacher (e.g., Ball & Bass, 2000; Francisco & Maher, 2005; Goos, 2004; Harel & Rabin, 2010; Rasmussen & Marrongelle, 2006; Yackel & Cobb, 1996). This instruction often involves teachers establishing social interaction patterns to set the expectation that students are expected to use reasoning, rather than authority, to gain mathematical conviction. For instance, Yackel and Cobb (1996) described an episode from a second grade classroom where a student correctly answers a question, but then changes her mind when the teacher repeatedly asks the class if her answer is correct. In response, the teacher repeatedly asks the student her name to which the student always replied Donna Walters. The teacher then asks if she would ever change her answer to Mary. When she replied no, the teacher said, "You know what your name is. If you're not sure you might have said your name is Mary. But… you know your name is what? Donna. I can't make you say your name is Mary. So you should have said, 'Mr. K, [the answer is] six. And I can prove it to you'. I've tried to teach you that" (p. 468-469). In this interaction, the teacher was attempting to make it normative that students base their answers on their own reasoning, not social cues from the teacher. In general, researchers have argued that students are more likely to rely on their own reasoning rather than authority in classroom environments where students are expected to justify their mathematical assertions, the students and not the teacher are the arbiter over whether a mathematical claim is correct or whether an argument is persuasive, and teachers attend not just to students' answers but also to their reasoning and argumentation (e.g. Francisco & Maher, 2005; Harel & Rabin, 2010).

Empirical Evidence

Schoenfeld (1983) noted that some of students' difficulties with problem solving could be explained by an empiricist views of mathematics, where students believed that mathematical claims were true or the procedures they used were valid because they appeared to work; this is in contrast to the mathematician in Schoenfeld's study who had what he coined was a rational[5] perspective and justified his claims and methods with logical deduction. Several studies document that students with empiricist views of mathematics were more likely to exhibit desirable metacognitive and problem-solving behaviors (Lester & Garafalo, 1987; Muis, 2008; Schoenfeld, 1983). It is widely accepted that empirical evidence is limited, and it is inappropriate to use this type of evidence to

[5] One with a "rational" perspective is one who relies on deduction to establish mathematical claims. To avoid misinterpretation, we are following Schoenfeld's use of the term and do not wish to judge those with an empiricist perspective to be "irrational."

gain certainty that a claim is true when one is in a mathematical context (e.g., Harel & Sowder, 1998; Recio & Godino, 2001)[6].

Researchers in mathematics education have tried to systematically determine what types of evidence that students used for justification tasks (for a more thorough review, see Weber, 2010). Knuth, Choppin, and Bieda (2009) presented approximately 400 middle school children with six justification tasks in number theory and found that students provided empirical justifications (i.e., they checked that these assertions were true for specific examples) 36% to 81% of the time, depending upon the task. Healy and Hoyles (2000) examined the responses of nearly 2,500 secondary students in the United Kingdom when they were asked to prove two conjectures in high school algebra. They found that 34% of the responses for a familiar conjecture and 43% for an unfamiliar conjecture were empirical arguments. It is noteworthy that in both the Knuth et al. and the Healy and Hoyles studies, the students were following curricula encouraging students to produce deductive arguments. Recio and Godino (2001) found a similar result at the university level. They asked over 400 first-year students to prove a statement from algebra and geometry and found that roughly 40% of the responses for these tasks were empirical. Such arguments appear to be less common amongst advanced mathematics majors. Iannone and Inglis (2010) asked 76 mathematics majors to attempt four proof production tasks, which yielded a total of 221 attempted proofs. Only two of these were empirical arguments. Weber (2010) conducted an interview study in which he presented

[6] Such a viewpoint is implicit in the Common Core Content Standards Initiative (2012) where students are expected to "make conjectures" and "make plausible arguments" by "reason[ing] inductively with data" but "build a logical progression of statements to explore the truth of these conjectures." Here inductive data is a tool for making conjectures but ought not be used to explore the truth of these conjectures.

28 mathematics majors who had completed an introductory proof course with an empirical argument and asked if they found it to be convincing or to be a proof[7]. The mathematics majors largely rejected the argument; only one student found the argument to be fully convincing and only two found it to be an acceptable proof. Further, 26 of the 28 participants explained their judgments by citing the limitations of empirical reasoning.

We find the methods used to investigate these issues to be somewhat problematic, essentially because we do not know the epistemic aims of the students as they formed their justifications. The experiments often did not make it clear whether the students were trying to obtain complete conviction that the mathematical assertion in question was true, a high level of conviction, or as much conviction was possible for them in the given situation. The students may also have had other non-epistemic aims such as obtaining partial credit or pleasing the interviewer (cf., Vinner, 1997). Knuth, Choppin, and Beida (2009) observed that their middle school students were more likely to provide empirical justifications for complicated statements than for simple ones and conjectured that students may have given empirical justifications because deductive justifications were too difficult. Healy and Hoyles (2000) noted that even though empirical justifications were common in their study, most students were aware of their limitations; an analogous result was obtained with pre-service teachers by A. Stylianides and G. Stylianides (2009). Thus we believe more research in this area with a more nuanced lens is needed.

Some researchers have designed instruction with the goal of changing students' epistemic beliefs about empirical evidence (e.g., Ball & Bass, 2000; Harel, 2001; Maher & Martino, 1996; Martino & Maher, 1999; G. Stylianides & A. Stylianides, 2009). For

[7] Note that this differed from the previous studies in that students were asked to determine the status of an existing argument, rather than producing one.

instance, researchers commonly ask students if the statement "$1141n^2 + 1$ is not a perfect square for all natural numbers n." After trying several cases, students are convinced that the claim is true. Indeed the claim is actually true for all natural numbers less than 10^{25}. However, there are very large counterexamples to this claim for integers greater than 10^{25}. Researchers expose students to this statement in the hopes that this experience will illustrate to students that one cannot be sure a mathematical assertion is always true from empirical evidence, even if one checks an extremely large number of cases. Ideally, this encourages students to search out new ways of verifying mathematical assertions (e.g., Harel & Sowder, 1998; G. Stylianides & A. Stylianides, 2009).

Deductive evidence

Most mathematics educators (and presumably most mathematicians) believe that a valid deductive argument (i.e., a mathematical proof) should be, in principle, sufficient to guarantee that a mathematical assertion is true. With this assumption, if an individual produces a proof of a statement or accepts that a proof of that statement is correct, this obviates the need to seek any other evidence in favor of that statement. However, this is not how students always behave. Fischbein (1982) presented a large number of high school students with a proof that $n^3 - n$ was always divisible by 6. Many students who accepted the proof as valid still desired to check the claim with a large number to be sure the claim was correct. Chazan (1993) illustrated how some high school students believed that a proof of a geometry theorem applied only to the specific diagram that was drawn but not all figures that satisfied the hypothesis of the theorem. He also reported that some students viewed proof merely as evidence in support of a statement, rather than a guarantee that the statement must be true. Harel and Sowder (1998) and Simpson (1995) independently remarked, in passing, that even after proving a theorem, undergraduates may still harbor some doubt that the theorem is true until they check it with examples; Harel and Sowder also claimed that some undergraduates might not be convinced by their own proof until an authoritarian source tells them that their proof is correct.

Theoretical perspective and sources of evidence for this paper

The arguments in this paper are based on two theoretical assumptions on our part. The first is that claims about how mathematicians gain conviction and how students ought to gain conviction should not be based on philosophical inquiries into how mathematical claims should be justified. Rather it should be based on an examination of what mathematicians actually do (Rota, 1991). Until recently, philosophers of mathematics have ignored mathematicians' practice. Cellucci (2009) claimed that in the last century the philosophy of mathematics was normative, dictating to mathematicians how they *ought* justify their claims; "mainstream" philosophers of mathematics are primarily concerned with how justification in mathematics should proceed within the analytic philosophy tradition. More recently, a growing number of philosophers have shifted their attention away from how mathematics should be justified and have begun studying how mathematics actually develops in the mathematical community. Cellucci (2009) refers to these philosophers as "dynamic," noting they view mathematics as a growing body of knowledge and are concerned with how mathematical practice contributes to this growth. An increasing number of philosophers are adopting the latter perspective and their research illustrates how many widespread beliefs about mathematical practice are inaccurate (e.g., Buldt, Lowe, and Muller, 2008; Davis & Hersh, 1981; Lakatos, 1978; Rota, 1991; Tymoczko, 1986). We use recent advances by these philosophers as well as a large number of empirical studies on mathematical practice that we have conducted to obtain a better understanding of mathematicians' practice. Recently, several mathematics educators have drawn from this philosophical literature to provide insights about the nature of proof (e.g., de Villiers, 1990; Hanna, 1991, 1995; Hanna & Barbreau, 2008; Harel & Sowder, 2007; Knuth, 2002; Mariotti, 2006; Otte, 1994). In this article, we contribute to this movement in three ways. First, we present a more comprehensive synthesis of the philosophical literature on this topic than others have reported. Second, we complement these philosophical analyses with the results of many empirical studies that we have conducted to gain insight into their mathematical practice. Third, we use the insights gained from this analysis to provide instructional recommendations and encourage further research to determine if these recommendations are viable.

The second construct that informs our paper is the notion of epistemic aims. Chinn, Buckland, and Samarapungavan (2011) defined epistemic aims as "goals related to finding things out, understanding them, and forming beliefs" (p. 146). They argued that although epistemic aims are usually are ignored by educational psychologists who study students' epistemic cognition, epistemic aims are essential to interpreting individuals' behavior. Individuals with different epistemic aims may engage in different activities and use different sources of knowledge to achieve these aims. The notion of epistemic aims is recognized as important for mathematics, both in terms of how mathematicians interact with one another, including seeking and conveying information (Aberdein, 2006), and in how the quality of a mathematical contribution is valued (Tao, 2007). Although there are a large number of distinguishable aims in mathematics, (Aberdein, 2006; Tao, 2007; see also Rav, 1999; Steiner, 1978), in this paper we focus on four relatively straightforward epistemic aims: (a) convincing themselves that a mathematical assertion is true, (b) forming a deductive argument that would be sanctioned by his or her peers (i.e., produce a proof) establishing that the assertion is true, (c) verifying that their own proof or the proof of a colleague is correct, and (d) increasing their understanding of the mathematics behind the mathematical assertion[8]. We believe distinguishing between these epistemic aims is especially important. A key point made in this paper concerns why we want students to use deductive evidence rather than empirical and authoritarian evidence. We argue that it is *not* the case that empirical or authoritarian evidence cannot yield rational conviction in mathematical assertions and that it is *not* the case that mathematicians do not rely on these sources of evidence to gain conviction. Rather it is the case that relying *exclusively* on these sources of evidence and not using deductive evidence denies students learning opportunities that can generate substantial mathematical understanding.

The three sections that follow focus on the role of authoritarian, empirical, and deductive evidence. In each section, we describe fundamental values that we believe nearly all mathematicians hold. We then use the philosophical literature to illustrate that mathematicians enact these values in different ways than some mathematics educators claim. Next we provide further support for these philosophical insights with empirical

[8] A full account of epistemic aim (d) is beyond the scope of this paper. By *understanding*, we mean three things: providing an explanation for why a theorem is true (i.e., highlighting a specific connection that makes it obvious to the reader why the theorem is true; this is often contrasted with longer technical proofs that do not provide this insight) (Steiner, 1976), learning or creating new methods to solve problems (Rav, 1999), discovering new relationships between mathematical concepts (Rota, 1991), and finding a new way to represent a mathematical concept (Thurston, 1994).

studies that we conducted. Finally, we use these insights to offer instructional suggestions.

Mathematicians' use of authoritarian evidence

Values in mathematics about authority

We believe that most mathematicians would agree with the following claims: If a mathematician wishes to advance a mathematical assertion as true, then it is incumbent upon that mathematician to present an argument that the claim is true. The argument that the mathematician presents should be *impersonal* and *transparent*-- that is, the argument should not be based on the mathematician's personal experience but rather should be sufficiently clear that other mathematician can check the arguments themselves. In other words, it is unacceptable to say "trust me" within such an argument. Anyone in the mathematical community has the right to challenge this argument and the mathematician proposing the argument has the obligation to defend it. If an argument is presented by an eminent mathematician and a flaw is found, that argument should be rejected.

We believe each of these values is productive. Some mathematics educators have designed mathematical classroom environments that supported these values, creating *communities of inquiry* and *cultures of argumentation* in mathematics classrooms (e.g., Goos, 2004; Maher & Martino, 1996; Yackel & Cobb, 1996). We fully support these aims and consider the design of these environments to contribute significantly to improving mathematics education. What we focus on in this section is what these values do not imply: In short, we argue that mathematicians frequently believe mathematical assertions are true on the testimony of others and that the perceived authority of the mathematician advancing a claim influences which testimony mathematicians choose to accept. Some mathematicians use authority to determine what arguments they choose to read, how they read these arguments, and how persuasive they find them.

Philosophical perspectives on authoritarian testimony in mathematics

We know much of what we know by what has been told to us by others (Coady, 1992). This is certainly the case with our everyday knowledge; for instance, most people routinely take medicine that their doctor prescribes for them without verifying that the medicine is safe and effective simply because most do not have the time and scientific expertise to verify these things for themselves[9]. Chinn, Buckland, and Samarapungavan (2011) contended that this is also the case with scientific knowledge. "Scientists' knowledge of their own areas of expertise derives largely from the testimony of their colleagues, because they read and hear about many more experiments than they can conduct themselves" (p. 153). The authors further argued that this is even the case with results from a researcher's own studies. We frequently trust that our research assistants used proper experimental techniques, accurately recorded the data, and did not fabricate or falsify data (cf., Lipton, 1988).

One could argue that mathematical knowledge is different in principle (e.g., Geist, Lowe, and Van Kerkhove, 2010; Weber & Mejia-Ramos, 2013a). If a scientist reads an experimental result in the literature, he or she may not have the equipment to replicate this study, and--even if the scientist did--he or she would not expect to obtain an identical result due to the vagaries of empirical sampling. However, if a mathematician reads the result of a certain calculation in a published paper, in principle that mathematician could

[9] Note that even if one uses medical websites to check that the medicine is safe, this is relying on the testimony of the author of that website.

perform that calculation himself or herself. One does not need to trust the author that he or she is correct. While perhaps true in principle, recent philosophical articles and reflections by mathematicians on their practice suggest that such personal verifications on published research articles are often not done (Auslander, 2008; De Millo, Lipton & Perlis, 1979; Geist, Lowe, & Van Kerkhove, 2010; Grcar, 2013).

It should be noted that it appears that *some* mathematicians actually personally verify every result that they use in their own research (or at least claim to do so). But many do not. Geist, Lowe, and Van Kerkhove (2010) noted:

> “we know a substantial number of mathematicians who want to understand all proofs that form a part of their papers and who will reprove even classical statements to be completely sure of their own results based on them; but we also know that many mathematicians are not as meticulous and accept results from the published literature as black boxes in their own research” (p. 158-159).

Auslander (2008) corroborated that mathematicians frequently trust theorems published in journals without checking their proofs, arguing that mathematics could not be a coherent discipline without mathematicians trusting the reviewing process.

One reason that mathematicians accept published results is to make efficient use of their time. The complete proofs of some mathematical theorems are extremely long. For instance, one of the major accomplishments of 20th century mathematics was the classification of finite simple groups. This classification spanned many decades and dozens of papers. It has been estimated that a complete proof of the classification of finite simple proofs would run over 10,000 pages. Most mathematicians accept that the problem of classifying the finite simple groups has been solved, yet few mathematicians have read a complete proof of this solution[10]. The preceding example is an extreme case, but it is not uncommon in mathematics for a useful result to be based on a lengthy proof that perhaps references several other papers containing lengthy proofs. If a mathematician wanted to use these types of results, then the mathematician would have no choice but to either invest a great deal of time verifying the results or accept them as true because they were published in the literature.

A second reason for relying on the correctness of published results is technical expertise. Mathematics is a fragmented discipline, and it is frequently the case that a mathematician studying one area of mathematics is unable to follow the proofs of another area of mathematics (e.g., Thurston, 1994). Consequently, many mathematicians do not have the background to verify many of the proofs of results that they read that are outside their areas of interest. In these cases, mathematicians often trust the experts and accept the proof as correct (Auslander, 2008).

In the preceding paragraphs, we claimed that mathematicians are willing to accept results as true because they underwent a peer review process. We further maintain that the reputation of the author influences how mathematicians review proofs that have not been published and that this happens in several different ways. First, mathematicians have to decide what unpublished proofs that they choose to read. As Aronson (2008) noted, if he read every preprint that was sent to him, he "wouldn't have time for anything else," so he is constantly weighing the risk of missing a significant mathematical advance against "spend[ing his] whole life proofreading the work of crackpots." He noted "in deciding whether to spend time on a paper, obviously the identity of a paper plays some

[10] For a philosophical discussion of such complex proofs, see Aschbacher, 2005.

role. If Razborov [an esteemed mathematician] says he proved a superlinear circuit bound for SAT, the claim on our attention is different than if Roofus McDoofus said the same thing."

Further, mathematicians treat difficulties that they encounter in a proof differently depending on who wrote the proof. In the mathematical community, there are some authorities who have a reputation for producing reliable work without errors. When another mathematician reads a part of an argument that seems to be dubious or confusing, that mathematician will be more likely to give the benefit of the doubt to the author, either by investing more time to try to figure out why the argument is correct or, in cases where the reader lacks the expertise to follow the argument, simply assuming that the argument is correct (Auslander, 2008; Muller-Hill, 2010). To take a recent high-profile example, Shinichi Mochizuki, a highly-regarded mathematician, announced a proof of the abc Conjecture (a well-known unsolved problem). The proof turned out to be extremely complex and hard to follow, and several months after Mochizuki publicly posted his work, it had still not been verified by colleagues. Rehmeyer (2013) commented that the proof is "so peculiar that mathematicians might have dismissed it as the work of a crank, except that Mochizuki is known as a deep thinker with a record of strong results."

Recent empirical studies

Recently, we conducted two empirical studies on how and why mathematicians read published proofs in journals. The first was an interview study with nine highly successful mathematicians from one of the top mathematics departments in the United States (Weber & Mejia-Ramos, 2011). The second was a survey sent to 118 research-active mathematicians (Mejia-Ramos & Weber, in press). In the interview study, we found one mathematician who refused to believe a theorem because it was sanctioned by other mathematicians. When asked why she read proofs that were published in journals, she responded, “I would like to find out whether their asserted result is true, or whether I should believe that it’s true. And that might help me, if it’s something I’d like to use, then knowing it’s true frees me up to use it. If I don’t follow their proof then I would be psychologically disabled from using it. Even if somebody that I respect immensely believes that it’s true” (Weber & Mejia-Ramos, p. 334). However, several mathematicians expressed the opposite viewpoint, stressing that they do not check these published proofs for correctness. For instance, one mathematician said, “Now notice what I did not say. I do not try and determine if a proof is correct. If it’s in a journal, then I assume it is” (p. 334). Another noted that while he has “certainly encountered mistakes in the literature, it’s not high on [his] mind” when reading a proof. Rather, he acts on the assumption that the proof is correct.

To test the generality of these findings, 118 mathematicians were shown the following three items on a survey and asked whether they agreed, disagreed, or were neutral to the statements using a five-point Likert scale:

C1: It is not uncommon for me to believe that a proof is correct because it is published in an academic journal.

C2: When I read a proof in a respected journal, it is not uncommon for the quality of the journal to increase my confidence that the proof is correct.

C3: When I read a proof in a respected journal, it is not uncommon for me to be very confident that the proof is correct because it was written by an authoritarian source that I trust.

We found that 72% of the participants agreed with C1 (12% disagreed), 67% agreed with C2 (17% disagreed), and 83% agreed with C3 (7% disagreed). This suggests that, for at least some mathematicians, that the source of a mathematical argument influences their judgment of it (Mejia-Ramos & Weber, in press). In summary, the majority of mathematicians claim that they will believe some proofs are correct because they were published in respected journals or written by an authoritarive source.

We also conducted studies on whether the author of an argument affects how mathematicians evaluate it. Inglis and Mejia-Ramos (2009) asked 190 research-active mathematicians to read a mathematical argument and rate how persuasive they found it on a scale of 0 through 100. Half the mathematicians were randomly assigned to the "anonymous group" and given the argument, but they were not told who wrote it. The remaining mathematicians were assigned to the "named group" and given the same argument, except these participants were notified that the argument had been taken from a textbook by J.E. Littlewood, a well-known mathematician. The named group's rating of the persuasiveness of the argument was more than 17 points higher than that of the anonymous group, a statistically significant difference (which was replicated in a subsequent study). One limitation to this study was that the argument discussed above was not a mathematical proof, although it was considered to be highly persuasive by the famous mathematician who produced it. Indeed, when mathematicians were shown deductive proofs in this study, the effect of authority vanished (although the authors conjectured that this may have been due to the relative simplicity of the proofs). Hence, it is plausible that all this study demonstrates is that mathematicians only consider the authority of the author when evaluating non-deductive arguments but not deductive ones. However, we have conducted another interview and survey study that suggests this is not the case.

In the interview study, Weber (2008) explored how eight mathematicians determined if a proof was correct. The interview was mostly task-based, where the participants were asked to evaluate the validity of eight purported proofs, but participants were also asked open-ended interview questions about how they determined the validity of a proof. One question asked if the author of the deductive argument that they were reading would influence how they read the argument. Six of the eight mathematicians claimed that it would. For instance, one mathematician responded:

> *There's something about trusting the source*. My assumption would be when reading a proof generated by a mathematician friend and they were pretty sure it was true, my assumption would be that the steps are correct and that I need to work hard to make sure that I can understand and justify every step. In a student proof, I'm less inclined to work through things that strike me as odd (Weber, 2008, p. 448; italics were from the original manuscript).

This finding corroborates the claims by Muller-Hill (2010) that mathematicians are less likely to dismiss a dubious argument if it came from an authoritative source. We tested the generality of this finding with a survey sent to 54 mathematicians who had experience refereeing mathematical papers. One survey item was:

When I referee a manuscript, it is not uncommon for me to be very confident that the proof is correct because it was written by an authoritative source that I trust. The response to this survey item was mixed; 39% of the surveyed mathematicians agreed with it while 41% disagreed with it (Mejia-Ramos & Weber, in press). These data do reveal that many, although not all, mathematicians claim to consider the reputation of the author of a proof when refereeing a manuscript.

Finally, we note that although many mathematicians assumed the published proofs they were reading were correct, they nonetheless frequently read proofs published in journals. Indeed, 74% of the 118 surveyed mathematicians agreed that they sometimes did not check a published proof for correctness, but read the proof to increase their understanding (Mejia-Ramos & Weber, in press). This corroborates a common claim in the philosophical literature: conviction is one purpose of proof for mathematicians, but expanding one's understanding is also a very important function of proof (e.g., Rav, 1999; Steiner, 1976).

Implications

In mathematics education, there are some authors who believe that students should not accept claims as true because an authority told them that this was the case and that one should not consider who wrote the argument while evaluating its validity. For instance, Harel and Sowder (2005) described the authoritarian proof scheme, i.e., seeking conviction from authority, as "an undesirable, yet common way of thinking" and instruction must attempt "to suppress the authoritarian proof scheme" (p. 42); Selden and Selden (2003) claimed, "mathematicians seem to treat proof as being independent from its author" (p. 6). Shanahan, Shanahan, and Misischia (2011) also hypothesized that mathematicians actively ignore the author of a mathematical argument and suggested that prospective teachers should be made aware that mathematicians read mathematics in this way. We believe that these suggestions are useful because they encourage students to take more autonomy in their learning, but that they are also oversimplified.

Some mathematics educators have emphasized the role of authoritarian evidence in the growth of mathematicians' knowledge. Notably, Hanna (1991) claimed the reputation of an author influenced whether a proof of a mathematical theorem was accepted, and Housman and Porter (2003) claimed that mathematicians sometimes use an authoritarian proof scheme in areas outside of their expertise. However, to our knowledge, researchers have yet to discuss the pedagogical relevance of these insights. Below we offer the following suggestions:

(i) Mathematicians often accept a published result as true without verifying the result for themselves. It should therefore not be surprising, or even disappointing, that students will accept a theorem as true because their teacher or textbook said it was so. To avoid misinterpretation, we emphasize that we strongly agree with mathematics educators that it is highly problematic if authorities are the *only* sources of students' mathematical knowledge, and we consider the design of mathematics classrooms that lead students to rely on their own resources to be valuable. However, as teachers and textbooks are usually correct, we claim that the students are not behaving irrationally or different from mathematicians if they believe what their teachers or textbooks tell them.

(ii) Some mathematicians will believe and use a result without checking its proof because they lack the technical expertise to do so. This finding is relevant for instructional practice. In many mathematics classes, there are claims of fundamental importance that we would like our students to know and use, even though these students do not yet have the sophistication to follow a demonstration that these claims are true. As a clear example, we expect high school students to know that π is irrational, but the proof would be challenging to most with a bachelors degree in mathematics. Also, calculus teachers require students to use the Fundamental Theorem of Calculus-- an indispensible tool for computing integrals-- even though they acknowledge that students are unable to understand why this theorem is true[11] (Weber, 2012). If we want students to know that π is irrational or to use the Fundamental Theorem of Calculus, we require them to rely on authoritarian evidence.

(iii) Some mathematicians will not check a published result for the sake of time. This finding may also be relevant for instructional practice. There are some important results in high school mathematics, such as the fact that the derivative of the sine function is the cosine function or the trigonometric addition formulas, whose justifications might be potentially comprehensible to secondary students. Nonetheless, the proofs of these statements are conceptually difficult for this population. In these cases, the instructor or curriculum designer needs to weigh the benefits that students would get from observing these justifications against the length of time it would take to present them, just as mathematicians make the same judgment when they choose which proofs to read.

(iv) Mathematicians are often willing to give the benefit of the doubt when an argument comes from a trustworthy source. We believe it would be useful for students to adopt the same stance when reading a proof that they find dubious that was given from a

[11] Here some mathematics educators might object that it would be possible to understand this theorem if calculus instruction was structured differently. We agree, but also acknowledge that most calculus teachers presently do not have the capability to do so. At present, most calculus instructors need to simply tell students the Fundamental Theorem of Calculus is true, present a proof that students are not capable of understanding, or not have students engage in integration.

trustworthy source. They should not immediately dismiss the proof by a teacher or textbook as wrong because they cannot understand it or it seems to contain a mistake. Rather, they should act on the assumption that the proof is likely correct and attempt to resolve their failure to understand it. To avoid misinterpretation, we are *not* saying that textbooks or teachers never make mistakes or that students should simply accept the proofs from each source as true. Students should of course have the right to ask their teacher to explain the proof if they are unable to resolve their disagreement with it.

(v) Finally, the issue of epistemic aims when reading proofs is important. When reading a published proof, mathematicians are often not reading the proof to check it for correctness or gain conviction in the theorem (Rav, 1999; Steiner, 1978; Weber & Mejia-Ramos, 2011). Rather, they are using it to gain mathematical understanding. (The importance of the multifaceted role of proof has been stressed by many mathematics educators, such as de Villiers (1990), Hanna (1990), Harel (2001), Healy and Hoyles (2000), and Knuth (2002), among others). This has an important implication for classroom instruction. The reason we present proofs to students or ask them to prove results themselves is *not* just so they can gain certainty that the result is true. They may already be convinced because the teacher told them that this was true. It is to expand their mathematical understanding. The largest problem with students relying predominantly on authorities to generate mathematical knowledge is *not* that students will come to believe things that are not true. It is that this practice is denying students many opportunities to generate mathematical understanding.

Mathematicians' use of empirical evidence

Values in mathematics about empirical evidence.

We believe that most mathematicians would agree with the following claims: Empirical evidence has significant limitations. There are mathematical claims that hold true for a large number of cases that are false in general. For this reason, verifying that a claim holds true for several examples, even an extremely large number of examples, is not *necessarily* sufficient for mathematicians to believe the claim is true. Further, empirical evidence is never sufficient for a claim to be regarded as proven, and mathematicians regard conjectures as open even if there is overwhelming empirical evidence in support of them.

We support instruction that is based on these values. Students should be aware that empirical arguments have limitations and are not an acceptable form of proof. Further, they should be critical in deciding the strength of empirical evidence in support of a mathematical claim (i.e., they should not believe a claim is true simply because they verified it for a small number of examples). However, we argue that it does *not* follow from these values that mathematicians are *never* convinced by empirical evidence and we challenge instruction that aims for students to never seek conviction in this way.

Philosophical perspectives on empirical evidence in mathematics

Empirical or inductive knowledge has been a central problem in the philosophy of science at least since David Hume. If a proposition has always been true in the past, even in a very large number of instances, what assurance do we have that it will continue to be true in the future? While empirical observations appear to be a necessary evidence for most scientific claims, they are, in principle, avoidable in justifying mathematical results[12]. A mathematician can prove that a claim about a mathematical concept is true by demonstrating that it is a necessary logical consequence of how that concept is defined.

Most published claims in mathematics are justified by mathematical proof; further, a claim is not sanctioned as a theorem by the mathematical community until it has been proven. Nonetheless Paseau (2011) stressed that this does *not* imply that mathematicians do not gain conviction via inductive evidence as well:

> "Mathematicians prefer deductive evidence and actively look for it even in the presence of overwhelming inductive evidence. The reason for this is that they are mathematicians and as such value deduction […] Given that the ends of mathematical activity are not solely to raise rational degree of belief in mathematical truths, it is accordingly not a cogent objection to our argument that mathematicians go on seeking proofs of [conjectures] in the presence of overwhelming inductive evidence. Compare marathon-running, whose ends are to get to a destination 24 miles away [sic] *and* to run there. If our sole end is to reach the final destination than any means will do; in particular, we may drive 24 miles by car. It is not a cogent objection to driving that if one had marathon-runners' ends one would shun the car in favor of running" (p. 144).

In short, to mathematicians, the epistemic aim of proof is not *only* to gain conviction (and some philosophers claim it is not even *primarily* to gain conviction—e.g., Cellucci, 2009; de Villiers, 1990; Rav, 1999). Proof satisfies other epistemic aims such as promoting

[12] There are arguably instances where this is not true, such as weighing the plausibility of a new axiom or an undecideable statement, but such a discussion is beyond the scope of this paper.

mathematical understanding[13], and mathematicians view proof as an end in and of itself. Hence, one cannot use the fact that mathematicians attempt to prove their assertions as a basis to claim that mathematicians do not obtain conviction via non-deductive arguments. In short, the act of proving has other epistemic aims beyond obtaining conviction.

Philosophers have used Goldbach's Conjecture to illustrate how mathematicians justifiably gain certainty in statements by empirical evidence. Goldbach's Conjecture asserts that every even number greater than 2 can be written as the sum of two primes. For instance, 4 can be expressed as 2 + 2, 6 can be expressed as 3 + 3, 8 can be expressed as 3 + 5, 10 can be expressed as 3 + 7 or 5 + 5, and so on. This conjecture is unproven and, indeed, mathematicians do not anticipate a proof in the foreseeable future (Baker, 2009). However, there is strong empirical evidence in favor of Goldbach's Conjecture. First, the conjecture has been empirically tested for an extremely large number of cases (4×10^{18}, at the time that this paper was written). Second, we can define a function $G(n)$ to express the number of distinct prime pairs that add up to n; for instance, $G(10) = 2$ because there are exactly two prime pairs that sum to 10, 3 + 7 and 5 + 5. $G(n)$ tends to increase as n increases, although not monotonically. For all numbers n greater than 100,000 that have been empirically tested, $G(n)$ is greater than 500. The notion that for some extremely large untested n that $G(n)$ would not only dip below 500 but have a value of 0 (which is what would be needed to disprove the Goldbach Conjecture) seems inconceivable. Most mathematicians believe this empirical evidence is overwhelming and regard Goldbach's Conjecture as true, even in the absence of a proof (e.g., Baker, 2009;

---

[13] In an influential mathematics education article, de Villiers (1990) listed other purposes of proof, such as systematizing a theory, discovering new theorems, and facilitating communication and debate.

Echeverria, 1996; Paseau, 2011). Echeverria claimed, "the certainty of mathematicians about the truth of GC [Goldbach's Conjecture] is complete" (p. 42).

While some philosophers of mathematics contend that there are instances where empirical evidence is sufficient to gain conviction, two important caveats should be noted. First, mathematicians gain this conviction based on a large number of examples. Indeed, Borwein (2008) noted that the development of powerful computing devices that can check a large number of cases greatly increased mathematicians' abilities to secure mathematical knowledge by inductive evidence. Second, mathematicians use their understanding and intuition about mathematics to determine when, and what types, of empirical evidence would be compelling. For instance, Baker (2009) argued that empirical evidence is arguably appropriate to gain full conviction in Goldbach's Conjecture because if a counterexample were found, it would likely be a rather small number; if the first counterexample to a mathematical claim was likely to be an extremely large number, empirical evidence would not be persuasive. For further examples of mathematicians being persuaded by empirical evidence and the conceptual considerations that enabled this conviction, see Putnam (1979, p. 55-57) who discussed why empirical evidence convinces mathematicians that the Twin Primes Conjecture is true, and du Sautoy (2004), who described a similar instance with the Reimann Hypthesis[14]. In short, it is not always the case that checking a mathematical proposition for an extremely large number of instances will provide compelling evidence that the proposition is true; there

[14] It is interesting to note that Baker (2009) argued that he was *not* convinced that the existing empirical evidence establishes that the Reimann-Zeta Hypothesis is true and Putnam (1979) did not find the empirical evidence in favor of Goldbach's Conjecture to be compelling. This illustrates an observation from our empirical studies—that there is substantial variance amongst mathematicians in how persuasive a specific collection of empirical evidence is.

must also be conceptual reasons for why empirical evidence is appropriate.

As a final note, de Villiers (1990) and Polya (1954) suggested that mathematicians are usually highly confident that a result is true before they attempt to prove it, with de Villiers (1990) noting that it is unlikely that mathematicians would invest the time in searching for a proof of a theorem if they did not believe the theorem was true.

Empirical studies

We have conducted several studies on how convincing mathematicians find empirical arguments. Weber (2013) investigated what types of empirical evidence that mathematicians found convincing. Weber's hypothesis was that mathematicians would find empirical evidence showing that a set of numbers has an uncommon property more persuasive than empirical evidence to showing that a set of numbers lacks an uncommon property. For instance, this hypothesis would predict that mathematicians would be more persuaded by empirical evidence showing that a set of numbers is prime than a set of numbers is composite as the latter is more likely to occur by chance.

To test this hypothesis, Weber (2013) conducted a study in which 97 mathematicians were randomly assigned to a Property or a Non-Property Group. The mathematicians in the Non-Property Group read an argument that all the terms in a particular sequence were never perfect squares. They were not explicitly told how this sequence was defined, but were informed that the sequence was "recursively defined using a small number of basic operations from arithmetic and combinatorics, specifically, addition, multiplication, and combinations." The first 12 terms of the sequence were listed, as well as their square roots. None of the square roots were integers, which clearly implies that the corresponding sequence terms were not perfect squares. Participants were also told that a computer program checked the first 10,000 terms of the sequence and no perfect squares were found. The mathematicians in the Property Group read a similar argument in support of the same claim. However, instead of seeing the square root of the first 12 terms of the sequence, they were shown that the first, third, fifth, and odd-numbered terms were all of the form $4k + 2$ and the even-numbered terms were all of the form $4k + 3$. They were also informed that a computer verified that this trend continued for the first 10,000 terms of the sequence. Since perfect squares are only of the form $4k$ or $4k + 1$, this implied that these terms of the sequence were not perfect squares. Participants in each group were asked how persuasive they found the argument on a scale of 0 through 100. They were also asked how persuasive the argument was using a multiple choice format, where the choices were (a) completely persuasive, (b) highly persuasive, as persuasive as I find some proofs that I read (especially if they are long and technical), (c) somewhat persuasive, and (d) not very persuasive (my confidence in the claim is not much higher as a result of reading this argument). Weber predicted that the mathematicians in the Property Group would find their argument more persuasive since being of the form $4k + 2$ or $4k + 3$ was an uncommon property (each happens only one quarter of the time), whereas not being a perfect square was a common property of numbers.

The results of Weber's study revealed three things about the way mathematicians regard empirical evidence. First, there was substantial variance in how the mathematicians in the Property Group evaluated their argument. Their mean persuasion rating was 44.0 but their responses ranged between 0 and 100. A histogram of participants' responses is given below. On their multiple choice responses, 1 mathematician indicated that he or she found the argument completely persuasive, 7 found it highly persuasive and as persuasive as some proofs that they read, 30 found it somewhat persuasive, and 11 found it not very persuasive. Hence, mathematicians show considerable variance in how they regard empirical evidence. Second, this evidence provides an existence proof that at least some mathematicians find some empirical evidence as convincing as some mathematical proofs that they read. Further evidence came from mathematicians' responses to a question at the end of the study. After their evaluations, all 97 participants were asked if "there were ever instances in their professional work in which they were convinced a mathematical claim was true because it had been verified with a large number of examples." The majority responded no, but 26 (27%) answered affirmatively. This illustrates that some mathematicians can find empirical evidence to be as persuasive as a proof or even completely persuasive. Third, as predicted, the Property Group found their argument to be more persuasive than the Non-Property Group. The Non-Property Group gave their argument a mean persuasion rating of only 16.0, significantly less than the 44.0 mean rating given by the Property Group; only two participants found the argument to be completely or highly persuasive and 36 of the 48 (75%) participants found the argument to be not very persuasive. This illustrates that mathematicians do not treat all empirical evidence equally. They use their conceptual understanding—in this case, how likely it was for numbers to not be perfect squares versus how likely it was to be of the form $4k + 2$ or $4k + 3$—to weight the strength of this evidence.

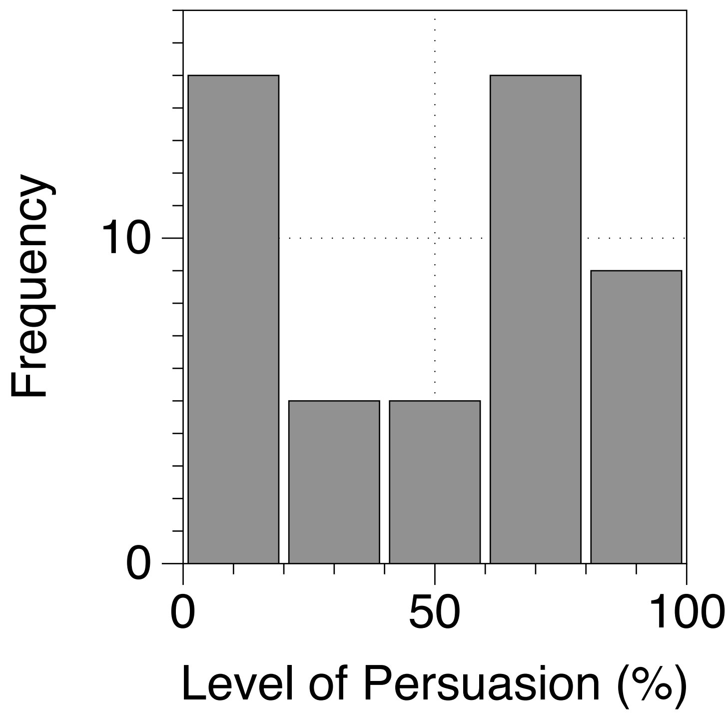

Figure 1. A histogram of mathematicians' persuasion ratings from the Property Group

A similar result was found in a separate study by Inglis and Mejia-Ramos (2009). They presented 180 mathematicians a more sophisticated empirical argument in support of the claim that the decimal expansion of π has, somewhere, one million 7's in a row[15]. The histogram of mathematicians' responses to this question is given in Figure 2 below. Figure 2 also reveals wide variance in mathematicians' evaluations. Some mathematicians found the argument to be highly persuasive or even completely persuasive (five participants rated the argument to be 100% persuasive) while others found the argument to be hardly persuasive at all (23 rated the argument to be 0% persuasive).

[15] This argument is based on the fact that empirical evidence suggests that most "ordinary" irrational numbers, including π, appear to have the mathematical property of being *normal*, which would imply that π would have one million 7's in a row at some point.

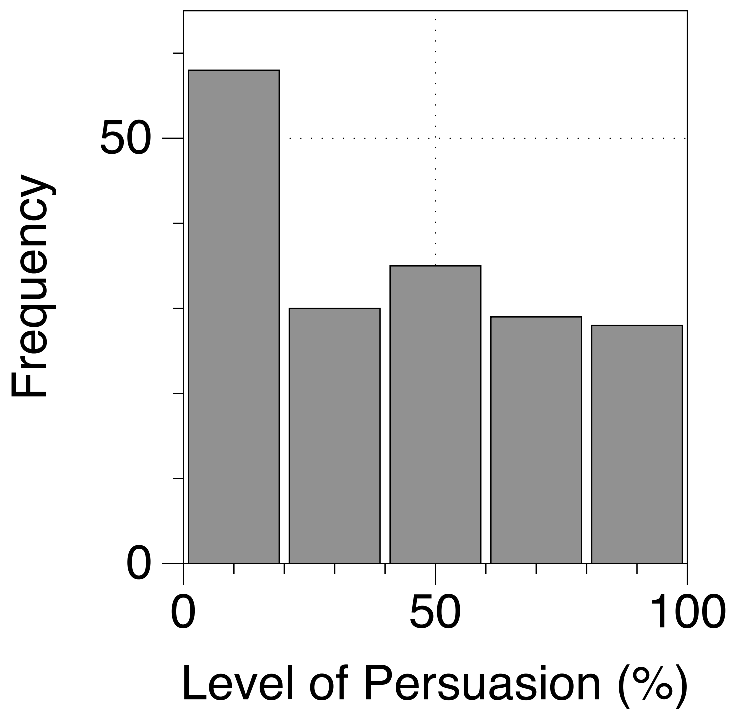

Figure 2: A histogram of mathematicians' levels of persuasion for an argument about the existence of a million consecutive 7s in the decimal expansion of π.

Implications

We contend the belief that mathematicians do not gain conviction from empirical evidence is inaccurate. Consequently, we believe the instructional goal of having students reject empirical evidence as a source of conviction—an explicit goal of several instructional programs (e.g., Harel, 2001; Stylianidies, 2007; G. Stylianides & A. Stylianides, 2009)—is problematic. We offer the following more nuanced suggestions:

(i) Empirical evidence is generally not recognized by mathematicians as an acceptable form of proof. Clearly students should be aware of this.

(ii) Nonetheless, there are instances where *some* mathematicians are convinced a claim is true on the basis of empirical evidence. These mathematicians still desire proofs of these claims, not for conviction but for other epistemic aims, including the mathematical understanding a proof might generate and because constructing proofs is a mathematical end in its own right. Thus, we believe that there are cases where students may gain legitimate conviction from empirical evidence. This is particularly relevant with dynamic geometry software, where both students and mathematicians regularly claim certainty in unproven geometric claims because the software allows the user to check these claims for a large number of cases (see de Villiers, 2004). It is not problematic that students are persuaded by this evidence, but students should know this evidence does not constitute proof and the quest for proof is still useful for the understanding a proof might generate.

(iii) Mathematics educators sometimes express concern that students believe things that are untrue based on empirical evidence (e.g., G. Stylianides & A. Stylianides, 2009). However, in these reported cases, students are usually basing their beliefs on a small number of empirical checks. In our view, the problem is not that students relied on empirical evidence; it is that they relied on *flimsy* empirical evidence. We agree that students should be dissuaded from accepting claims on such evidence, but believe students should be made aware of when it is appropriate to draw conclusions from empirical evidence, as well as what types of evidence and what amounts are compelling and why, rather than trying to persuade students to abandon accepting claims on empirical evidence altogether.

(iv) Mathematics educators commonly cite the claim “$1141n^2 + 1$ is never a perfect square for all natural numbers n” as a paradigmatic reason for why students should not trust empirical evidence. The claim is false in general, but the first counterexample does not occur until *n* is greater than 10^{25}. Given the analysis above, this is not a very good reason for abandoning empirical evidence entirely. Mathematicians are skeptical of empirical evidence showing that a set of numbers lacks an unusual property, such as being a perfect square, but more accepting of other types of evidence. Rather that contrast the limitations of empirical evidence with the certainty of deductive evidence, we suggest it might be more worthwhile to develop curricula to show how the two can work in tandem.

Deductive evidence

Values in mathematics about deductive evidence and proof

We believe that most mathematicians would agree with the following claims: Deductive evidence and proof are extremely important, both in the culture and history of mathematics and in the practice of contemporary mathematicians. One of the benefits of proof is that proof offers a *conditional guarantee* that a theorem is true (Fallis, 1997). If the reasoning in the proof is valid and the conceptual system in which the proof is couched is sound[16], the theorem is necessarily true. No other types of evidence can offer this type of conditional guarantee. Proof is also particularly valued for its *generality*. Proofs are of finite length but nonetheless offer a conditional guarantee that an infinite number of assertions are true. (e.g., the claim that every odd integer squared is odd

[16] The phrase "conceptual system is sound" masks a deep and difficult philosophical issue that is beyond the scope of this paper. The main point here is that a theorem can be false even it is supported by a proof that is logically correct. This can happen, for instance, if the logical theory in which the proof is couched is mathematically inconsistent (i.e., one can prove both a statement and its negation in the theory), the definitions of concepts do not accurately capture the meanings of those concepts (a point discussed in detail in Lakatos, 1976), or the assumed background knowledge of true results and sound methods of inference is unreliable.

implies “1 squared is odd,” “3 squared is odd,” and so on). Proof alone is how a claim can be sanctioned by mathematicians as a theorem. One feature of proof is that there are certain methods of inference that are clearly invalid and would be rejected by all (or nearly all) mathematicians. For instance, even though the theorem “all differentiable functions are continuous” is true, it would be unacceptable to use this theorem to infer that a function *f(x)* was differentiable because it was continuous. This is because “if A, then B” and “B” is not logically sufficient to conclude “A.”

We fully support instruction that is based on these values. Students should appreciate the importance of proof in mathematics culture and practice; they should value the high levels of conviction that proofs can provide and understand the generality of proof. Students should also recognize some forms of reasoning are invalid and unacceptable within a proof. The points that we wish to emphasize here are that a purported proof only offers a *conditional guarantee*, not an *absolute guarantee*, that a mathematical claim is true, proof has epistemic aims beyond providing conviction, and mathematicians do not always agree on what methods of inference are valid within a proof.

Philosophical perspectives on deductive evidence in mathematics

In the United States, much of a non-mathematician’s conceptions of proof are formed in their high school geometry course, as this is typically the only course in which most students have extensive experience with proof (e.g., Moore, 1994; Stylianides, 2007). In these courses, proofs are typically written in a two-column format (see Herbst, 2002), where every statement in a proof is something known to be true (e.g., an axiom, a definition, an assumption from the theorem being proven) or is a direct deduction from earlier assertions in the proof using a rule of inference that is explicitly stated. In the two-column format, the statements within the proof are given on the left side and how these statements were deduced is given on the right side. This format for proof codifies what many believe is the essence of proof—each statement in the proof is a direct logical consequence from previous statements. Using this format, we can see how a proof guarantees the truth of a theorem, and it is relatively unproblematic to check that a proof is correct. Recent developments in the philosophy of mathematics, as well as empirical studies, challenge these claims about proof.

In mathematical practice, the proofs that mathematicians write are not nearly as structured or detailed as the typical two-column geometry proof. Most proofs have substantial gaps, where it is not immediately obvious why some statements in these proofs are necessary logical consequences of previous statements. Sometimes it is claimed that these gaps are merely omissions of details that a knowledgeable reader could easily fill in, but Fallis (2003) denied that all gaps could be explained in this way. Instead Fallis claimed that some gaps in the proof were bridged neither by the author of the proof nor any reader of the proof. This implies that verifying a proof does not merely involve checking that specific explicitly stated rules of inference were correctly applied[17].

As verifying a proof is not mechanical, this opens up the possibility that a mathematician might mistakenly believe a step within a proof is correct because he or she failed to detect an error; consequently, the mathematician might accept an invalid

[17] As a dramatic example, Aschbacher (2005) noted that a full proof of the classification of finite simple groups would run over 10,000 pages and would, with probability 1, contain deductive errors that have not yet been detected, but nonetheless he considers the theorem to be proven. This implies that the conviction that the proof was valid cannot be entirely attributed to checking that each deduction in the proof was valid.

argument as a proof. De Milo, Liptus, and Perlis (1979) argued that even if we are reasonably confident that each individual step in a proof is correct, as proofs become increasingly lengthy, the possibility that *none* of the steps contains an undetected error becomes non-trivial. As a result, we cannot be sure that the proof, and therefore the theorem that it purports to establish, is correct. This concern was noted as far back as 1739 by David Hume, who noted:

> "There is no Algebraist or Mathematician so expert in his science, as to place entire confidence in any truth immediately upon his [sic] discovery, or regard it as anything, but a mere probability. Every time he runs over his proof, his confidence increases; but still more by the approbation of his friends; and is raised to its utmost perfection by the universal assent and applauses of the learned world" (quoted by Nathanson, 2009, p. 8).

Note that Hume is arguing that the certainty that one obtains in a theorem is not the result of producing a proof of it, but also by having friends and scholars check the proof for correctness. Paseau (2011) expressed a similar argument as follows: "That we are in possession of a proof of *p* does not imply we should be certain of *p* […] The proof may be long and hard to follow, so that any flesh-and-blood mathematician should assign a non-zero probability to its being invalid. The longer and more complex the proof, the less secure its conclusions" (p. 143). Inspection of proofs published by actual mathematicians reveals that this is not merely a philosophical concern. Davis (1972) estimated that half of published proofs contain errors, and Devlin (2003) discussed recent examples of proofs of important theorems that were thought to be valid by the mathematical community before a flaw was discovered. In an empirical study of the number of corrections published in various disciplines, Grcar (2013) found that mathematics papers were not significantly less likely to contain errors than papers from other disciplines. Due to findings such as these, Nathanson (2008) expressed a stronger concern that the mathematics literature was unreliable. In short, although deductive evidence is an extremely powerful source of justification in mathematics, a given deductive proof is often insufficient to gain full certainty in the truth of an assertion. Nathanson (2008) and Grcar (2013) both noted that mathematicians' refereeing practices, including sometimes checking the proof in terms of its high-level ideas but not its details[18], may contribute to the proliferation of errors.

A widely held belief about proof is that mathematicians have an unusually high level of agreement as to whether a deductive argument constitutes a valid proof[19]. For instance, McKnight, Magid, Murphy, and McKnight (2000) asserted that "all agree that something is either a proof or it isn't and what makes it a proof is that every assertion in it is correct." (p.1). Selden and Selden (2003) remarked on "the unusual degree of agreement about the correctness of arguments and the truth of theorems arising from the validation process" (p.7); they also claimed that validity is a function only of the argument and not of the reader: "Mathematicians say that an argument proves a theorem, not that it proves it for Smith and possibly not for Jones" (p. 11). It is not only mathematics educators who make this claim. The philosopher Azzouni (2004) attempted to explain why "mathematicians are so good at agreeing with one another on whether

[18] Whether mathematicians actually read a proof for its high-level ideas is currently under debate in the mathematics education community (e.g., Inglis & Alcock, 2012, 2013; Weber & Mejia-Ramos, 2013b).

[19] Note that we are not talking about alleged proofs whose validity is considered problematic by some members of the mathematical community, such as probabilistic proofs, picture proofs, or computer assisted proofs (cf., Aberdein, 2009).

some proof convincingly establishes a theorem" (p. 84). However, other philosophers have recently challenged the view, arguing that due to the wide range of mathematical practices, the notion that there is a single criterion on which proof is judged is implausible (e.g., Auslander, 2008; Rav, 2007).

Empirical studies.

We have conducted several studies that challenge the notion that mathematicians agree on what constitutes a proof and that their processes for evaluating a proof are sufficient to obtain certainty in its correctness. Our investigations into mathematicians' validation of proofs stemmed from an influential article that Selden and Selden (2003) published on mathematics majors' proof validations. In their study, Selden and Selden, both published research mathematicians, asked students to evaluate an argument from elementary number theory they labeled "the real thing," which they presented as a transparently valid proof, and another they labeled "the gap," which they evaluated as invalid. In a pair of independent studies, Weber (2008) and Inglis and Alcock (2012) presented these proofs to 20 mathematicians (8 in Weber's study and 12 in Inglis and Alcock's) and asked them to determine whether or not they were valid. Collectively, only 14 of the 20 mathematicians judged "the real thing" to be valid and only 12 judged "the gap" to be invalid. As these proofs were short, came from an elementary mathematical domain, and were chosen by former research mathematicians because of the ostensible transparency in their validity, we considered it highly surprising that they generated disagreement among the 20 mathematicians in our studies.

To explore this issue further, Inglis, Mejia-Ramos, Weber, and Alcock (2013) presented 109 mathematicians with an argument purporting to prove that

$\int x^{-1}dx = \ln(x) + c$. They were told that this argument was submitted to the expository mathematics journal, the *Mathematical Gazette*, and were asked to judge whether or not the argument was valid. The argument established this theorem by commuting the limit and integral sign, a mathematical technique that is not valid in general. The technique was applicable in this particular context, but the justification for why it was applicable in this context was absent from the proof. When asked if the argument was valid, 29 mathematicians (27%) agreed that it was and 80 (73%) said it was not, with applied mathematicians significantly more likely than pure mathematicians to judge the argument to be valid (50% vs. 17%). To ensure that the differences in participants' judgment were not merely due to performance error, we asked a follow-up question. Participants were informed that a mathematician critiqued this argument by noting that the limit and integral sign were commuted; they were asked if such a criticism was reasonable and if this alone was enough to render the argument invalid. For both groups, over 80% of the participants found the critique to be reasonable. Of the 29 who judged the argument to be valid, 24 (83%) claimed this critique was not sufficient to render the argument invalid. Only 24% of those who initially judged the argument to be invalid made this judgment. This suggests that mathematicians do not agree about whether particular inferences within an argument are permissible without justification, even in a domain as basic as elementary calculus.

In a separate line of studies, we explored how mathematicians checked the validity of an argument. Weber (2008) presented eight mathematicians with six to eight arguments and asked them to determine if these proofs were correct. In reading these arguments, there were 77 instances in which participants were not sure if a particular assertion within the proof was correct. In 15 instances, participants constructed a sub-proof verifying the assertion, arguably using deductive evidence to make certain the assertion was true. However, it was more common for participants to give informal arguments (33 instances) where general approaches to proving the assertion were stated but significant details were not carried out or with empirical evidence (19 instances) in which they verified that the assertions were true by checking them with specific examples. This empirical evidence usually consisted of looking at only a small number of examples, sometimes even a single example, so these empirical checks most likely did not provide the mathematicians with complete certainty that the assertion they were verifying was necessarily correct. To test the generality of these observations, Mejia-Ramos and Weber (in press) asked 54 mathematicians who had experience refereeing papers if they agreed, disagreed, or were neutral to the following statement:

> When I referee a paper and I am not immediately sure that a statement in the proof is true, it is not uncommon for me to gain a sufficiently high level of confidence in the statement by checking it with one or more carefully chosen examples to assume the claim is correct and continue reading the proof.

Forty three percent of the surveyed mathematicians agreed with the statement, while 28% disagreed. This illustrates, by mathematicians' self-report, that some are not using a validation process to guarantee the correctness of a proof with absolute certainty.

In general, based on our interviews with mathematicians, mathematicians highly value empirical evidence in making sense of, and gaining confidence in, the proofs that they read (Weber & Mejia-Ramos, 2011). As one mathematician described to us, he "always" considered examples when checking proofs: "I always use examples to make sure the theorem makes sense and the proof works. I'm sure there are some mathematicians that can work at an abstract level and never consider examples, but I'm not one of them. When I'm looking through a proof, I can go off-track or believe some things that are not true. I always use examples to see that makes sense" (p. 338). When we checked the generality of this finding with the survey for the 54 mathematicians who had experience refereeing, we asked if they agreed, disagreed, or were neutral to the following statement:

> When I referee a proof, it is not uncommon for me to see how the steps in the proof apply to a specific example. This increases my confidence that the proof is correct.

Most surveyed participants (74%) agreed and only two of the 54 mathematicians (4%) disagreed. To these mathematicians, making sense of and checking deductive evidence was done, at least in part, by considering empirical evidence.

Implications

Based on our arguments above, we believe that some mathematicians do not gain complete conviction from some of the deductive arguments that they read. Although some mathematics educators have arrived at a similar conclusion (e.g., Hanna, 1991; Otte, 1994), others have argued that mathematicians do gain complete conviction via proof (e.g., Harel & Sowder, 1998). We do not believe that we should expect students to gain complete conviction from the proofs that they produce but instead offer the following more nuanced suggestions:

(i) Deductive evidence can be *highly* persuasive evidence in favor of an assertion. This is also the type of evidence that mathematicians use to sanction theorems. Students should be led to appreciate the power and generality of a proof and its role in establishing theorems. That students think of proof as merely evidence in favor of an assertion or a general proof as applying to a special case (cf., Chazan, 1993) is indeed problematic.

(ii) However, proofs, in general, do not provide mathematicians with *absolute* certainty that a theorem is true because often mathematicians cannot be absolutely certain that the proof is correct. With students who frequently struggle mastering the nuances of proof, we would expect proof to provide them with even less certainty. Consequently we do not consider it irrational or problematic that students desire further evidence in favor of a statement that they have proven. Asking an authority if their proof is correct or continuing to check the proven statements with examples does not seem to be irrational; indeed, we consider this consistent with mathematicians' practice.

(iii) Unlike the two-column proofs that students observe in some high school geometry courses, the proofs that mathematicians produce have gaps. Indeed, the two-column proofs that students counter in geometry are arguably not even idealized versions of the proofs that mathematicians produce (e.g., Baker, 2009; Fallis, 2003).

(iv) Whether a deductive argument is a valid proof is (sometimes) not an objective feature of the argument, but depends on the mathematician reading it, among other factors. While some arguments are clearly invalid (e.g., arguments based on blatantly false assertions or containing computational errors), other arguments' validity is more problematic to determine because one cannot be sure how large of a gap in a proof is permissible. This has an important implication for the teaching of mathematics majors; mathematicians often do not agree on what level of detail should be included in proofs they present to students (Lai, Weber, & Mejia-Ramos, 2012) and, we hypothesize, on the level of detail that students should provide on proofs that they submit for credit.

(v) In mathematics education, deductivist and empiricist epistemological beliefs are often placed in opposition to one another (e.g., Muis, 2004, 2008). Students are expected to abandon the empiricist beliefs that they frequently hold in favor of the deductive beliefs valued by mathematicians. We find this to be a conceptual error. Many mathematicians seek conviction in mathematical statements from deductive evidence *and* empirical evidence, and indeed, gain the most when these two sources of evidence are used in tandem. It is telling that in Muis' (2008) study of students' epistemological beliefs, students who held empirical and deductivist beliefs exhibited the same positive problem-solving behaviors as those who only evinced deductivist beliefs. The moral here is that mathematics students should develop an appreciation for the power, utility, and appropriateness of deductive evidence, but this does not entail completely abandoning empirical evidence.

Summary

Implications for mathematics education

In mathematics education, with respect to justification and proof, it is widely believed that the goals for instruction should be for students to have the same beliefs and justification practices as mathematicians. We agree with this premise, but believe that for this to be feasible, we need to have an accurate understanding of what mathematicians' practice is and we are concerned that current instructional goals are based on a somewhat shallow view of what mathematicians do.

In this paper, we set out to establish four claims about mathematical practice and to discuss their implications for the goals of mathematics instruction. These claims were: (1) mathematicians sometimes gain high levels of conviction from authoritarian sources or empirical evidence; (2) the sources that mathematicians consider depend upon their epistemic aims; (3) the quality of authoritarian and empirical evidence influences its persuasive power; and (4) there is substantial heterogeneity in how persuasive mathematicians find different types of evidence.

Regarding claim (1), we have argued that many mathematicians, by necessity, frequently accept mathematical results because they have been published in respected journals and gain confidence in these results because they were written by authoritarian sources. We have demonstrated that there were instances in which some mathematicians gained very high levels, or even complete certainty, in mathematical statements based on empirical evidence. Finally, we have argued that there are instances in which proofs do not and cannot provide mathematicians with complete certainty. Ergo, the instructional goal of having students gain absolute certainty from deductive evidence, and deductive evidence alone, is not consistent with mathematical practice.

Regarding claim (2), mathematical proof is valued by mathematicians, both for the mathematical understanding that it provides, as an end in and of itself, and the conditional guarantee that it can provide (i.e., if the proof is correct and the supporting theory is sound, the theorem must be true). Empirical evidence and authoritarian sources cannot be used to sanction a theorem in the absence of a proof and do not provide the gains in understanding that a proof can provide. However, when it comes to gaining conviction that a theorem is true or checking the validity of an argument, empirical evidence and authoritarian sources are frequently used by mathematicians. This claim has an especially important consequence for instruction. Deductive evidence should not be valued over empirical or authoritarian evidence because deductive evidence can provide certainty in mathematical statements while the other types of evidence cannot. Rather, deductive evidence should be valued because it can provide mathematical understanding that the other types of evidence cannot and because it is the only means that mathematicians use to sanction theorems in their professional practice.

Regarding (3), we have shown that mathematicians consider the quality of a journal in estimating the truth of its theorems. We have also shown that some types of empirical evidence (e.g., showing some members of an infinite set all have an unusual property) is more persuasive than other types of evidence (e.g., showing some members of an infinite set all lack an unusual property). Blanket claims about the persuasive power of authoritarian or empirical evidence are difficult to make.

Regarding (4), we have acknowledged that there are core values about conviction, justification, and proof that we believe are shared by most mathematicians. However, we have also shown that mathematicians disagree both on how persuasive they found specific empirical arguments and on whether a specific proof from calculus is valid. We have also found that the majority of mathematicians state that they rely on authority in accepting a published theorem as true but some mathematicians claim not to do this.

The view that students should gain conviction from deductive evidence, but not empirical or authoritarian evidence, is in some respects a very worthy one for mathematics instruction. An exclusive reliance on authority severely limits students' autonomy. A failure to appreciate the potential weaknesses of empirical reasoning will lead students to believe some mathematical statements that are not true. An inability to appreciate or understand deductive argumentation will deny students a valuable source of mathematical understanding and prevent them from participating in authentic mathematical discourse. Consequently, in our view, the instructional recommendations of focusing on deduction and limiting appeals to authoritarian or empirical evidence has a great deal of merit. However, as we argue above, we think the basis for these instructional recommendations is oversimplified and not based on an accurate understanding of mathematical practice.

For each of the main claims that we posit in this paper, we have outlined instructional recommendations that move beyond the view that students should gain complete conviction from deductive evidence and deductive evidence alone. We believe that these instructional recommendations are not only more consistent with mathematicians' practice, but also are more realistic and provide greater learning opportunities for students. Some of these recommendations concern what the goals of mathematics instruction should be. If we accept the assumption that students should develop epistemic beliefs and actions that are consistent with the practice of mathematicians, the theoretical and empirical arguments that we present in this paper provide strong support for these recommendations. Other recommendations are based on classroom practices that we would like students to engage in to develop their mathematical understanding. Clearly these pedagogical suggestions should be tested. One test could be seeing if classrooms that follow our recommendations (e.g., classrooms where some, but not all, claims are justified and it is acceptable for a student to check with a teacher to see if his or her justification is correct) lead to better or worse achievement than classrooms following recommendations common in the mathematics literature (e.g., classrooms where all claims are justified and students cannot check their work with an authority).

Implications for research on epistemic cognition

We believe that our research makes a number of claims that are important for research on epistemic cognition.

Experts rely on the testimony from authoritative sources. We have shown that some mathematicians frequently do this to decide which claims they accept and build upon in their work. We concur with Chinn, Buckland, and Samarapungavan (2011) and Bromme, Kienheus, and Porsch (2009) that it is simply inaccurate to claim, as some models of epistemic cognition do, that expert mathematicians and scientists always rely on their own reasoning to decide what is true. With mathematicians, this is even when they have the capacity to make this decision themselves.

Developing understanding is an important epistemic aim for expert practitioners. We have shown that mathematicians are often willing to accept a claim as true and a proof as correct if it was written by an authoritative source or published in a respectable outlet. Nonetheless, most mathematicians study these proofs. This is often not to increase their conviction in the correctness of the claim or the validity of the proof. Instead, they read the proof to increase their mathematical understanding.

Evaluating epistemic cognition requires understanding individuals' epistemic aims. We have argued that the epistemic actions that mathematicians take (and that students should take) are dependent upon their epistemic aims, which as Chinn, Buckland, and Samarapungavan (2011) argued was not present in many models of epistemic cognition. Empirical evidence and appeals to authority are often useful for the epistemic aim of finding out what is true, but less useful in mathematics for sanctioning claims or expanding one's understanding.

It is important to test our beliefs about experts' epistemic cognition with empirical studies. We discussed claims that some mathematics educators made about mathematicians' epistemic cognition that we found to be inaccurate and oversimplified. It is important to emphasize that the claims that we challenged were not generated randomly or haphazardly by these researchers; rather these were based on what these researchers felt was substantial evidence, including the writing of individual mathematicians as well as their own experience doing mathematics. Of course, this type of evidence is both appropriate and necessary for forming hypotheses about mathematical practice. The empirical studies we present in this paper caution us that our claims about experts' mathematical cognition, claims that may appear intuitively obvious, are sometimes wrong. Before using beliefs about experts' epistemic cognition as a basis for instructional design, we contend that it is important to conduct empirical studies to increase our confidence that these beliefs are correct. However, such empirical studies are notably absent from domain-specific research on epistemic cognition. Our claim is based on a simple premise: If we want to make claims about how mathematicians or scientists behave, what they claim, and what they value, it is useful both to actually talk to a large number of mathematicians and scientists and to systematically observe their behavior.

An alternative framework for models of epistemic cognition.

In an influential paper, Smith et al (1993) reconceptualized how the science education community viewed students' misconceptions about scientific phenomena. The traditional view was that there were normatively correct ways of understanding scientific phenomena and common incorrect ways that students frequently used to understand these same phenomena. These incorrect ways were labeled as misconceptions. The goal of instruction was to have students abandon these misconceptions and adopt the normatively correct ways of understandings. Smith et al (1993) challenged this viewpoint and offered an alternative perspective. The so-called misconceptions that students held were usually perfectly viable in students' previous experience. The problem was not that students believed such things but that they were applying them in a domain where they did not apply. For instance, the belief that "one needs to continually exert force on an object to keep it in motion" is not wrong. Indeed, it applies in nearly every real-world situation that a student will encounter. If you are driving and take your foot off the gas, your car will eventually stop. This belief is only inaccurate when working in a frictionless environment. The goal of science education should not be to eradicate misconceptions but rather to have students appreciate the scope of their viable conceptions and empower students by developing new conceptions that can be used to answer different types of questions.

We believe an analogy can be drawn with most traditional models of epistemic cognition. In these models, there are primitive naïve beliefs that students frequently accept and more sophisticated beliefs that are held by scientists and other enlightened individuals. The goal of education, according to these models, is to lead students to reject the common naïve beliefs and adopt the more sophisticated ones. For example, students should reject the notion that knowledge is handed down by authorities and accept that knowledge is generated by the individual. Students should not accept mathematical results because they appear to work but because they can be logically deduced. We believe this is a conceptual error. Many purportedly naïve epistemic beliefs and epistemic actions are both natural and correct *for some epistemic aims* and indeed are held and used by expert practitioners. The goal of instruction should *not* be to have students reject naïve epistemic beliefs; indeed, the notion that students should *never* believe something because an authoritative source told them that it was true is utterly infeasible for practical life. Rather, students should understand *how to implement these epistemic actions effectively* (for instance, how does one determine what constitutes reliable testimony from an authoritative source?) and for *what epistemic aims these actions are useful*. Further, students should *expand their repertoire of epistemic actions* to include those that are valued by researchers in epistemic cognition.

At a broader level, we find that educators often desire that students participate in some desirable behavior or hold some desirable belief (call it X) but instead observe that students predominantly engage in some other type of behavior or hold some other belief (call it Y). While we agree that the absence of X in students' reasoning may be a problem, it does not necessarily follow that the presence of Y is as well. That is, the behaviors that students *do* engage in and the beliefs they do hold are not necessarily problematic. This can explain what an apparent inconsistency between the literature on epistemic cognition and the results of our paper. For instance, Muis (2004) wrote that students "believe mathematical knowledge is passively handed to them by some authority figure, typically a teacher or textbook author, and that they are incapable of learning mathematics through logic and reason" (p. 330) and she observed that students holding these beliefs tended to have lower achievement than those who believed they could learn mathematics through logic and reason. Yet in this article, we argue that mathematicians do and students should rely on the testimony of authority figures in some situations. The quoted excerpt is a compound sentence containing two assertions: (i) students do not believe they can use deduction and reason to learn mathematics and (ii) students believe mathematical knowledge can be handed to them by an authority, with the implication that (i) and (ii) are both problematic. We agree that (i) *is a significant problem*. However, we think it is a fallacy to infer that (ii) is problematic as well. It is our hypothesis that the correlation with lower achievement reported above is not the presence of the belief that the testimony of authorities is a useful source of knowledge, but the absence of the belief that using one's own reasoning is also a valuable source of knowledge.

As we noted in this paper, Muis (2008) reported evidence that directly supports our hypothesis when describing students' empiricist and deductivist views of mathematics. Students who held strict empiricist views (i.e., the students primarily judged mathematical truth on what looks right or works) had lower achievement than students with strict deductivist views (i.e., the students primarily judged truth through mathematical logic). However, students who held a hybrid of empiricist and deductivist views were no worse off than those with strict deductivist views. In this case, the problem with holding strict empiricist views is not the presence of empirical beliefs, but the absence of appropriate beliefs about the value of logical deduction.

A consequence of this argument is that measuring individuals' epistemic cognition along a single continuum might not be viable. Indeed, as Greene, Torney-Purta, and Azevedo (2010) argued that if "students consider many different kinds of warrants (e.g., appeals to authority, logic, coherence) when justifying logic claims, attempts to quantitatively measure those beliefs about knowing on a single continuum are unlikely to succeed. Restricting justification to a single factor may be one reason that numerous studies of personal epistemology measures have found little evidence of construct validity and poor data-model fit" (p. 236). We believe more nuanced models considering how different types of evidence are used for different epistemic aims may provide a more accurate and coherent account of students' personal epistemology in mathematics.

Acknowledgments

Some of the research reported in this manuscript was supported by a grant from the National Science Foundation (DRL-0643734). We would like to thank the anonymous reviewers and the editor for helpful comments on revising this manuscript. All opinions expressed here are our own and not necessarily those of the National Science Foundation.

References

Aberdein, A. (2006). The informal logic of mathematical proof. In R. Hersh (Ed.), *18 Unconventional Essays on the Nature of Mathematics*. New York: Springer, pp. 56–70.

Aberdein, A. (2009). Mathematics and argumentation. *Foundations of Science*, 14, 1-8.

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

Aschbacher, M. (2005). Highly complex proofs and implications of such proofs. *Philosophical Transactions of the Royal Society A*, 363, 2401-2406.

Aronson, S. (2008). Ten Signs a Claimed Mathematical Breakthrough is Wrong. Downloaded from: http://www.scottaaronson.com/blog/?p=304. Last downloaded: April 8, 2013.

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

Azzouni, J. (2004). The derivation-indicator view of mathematical practice. *Philosophia Mathematica*, *12*, 81-106.

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

Baker, A. (2009). Non-deductive methods in mathematics. Contribution to the *Stanford Encyclopedia of Philosophy*. Downloaded from: http://plato.stanford.edu/entries/mathematics-nondeductive/. Last downloaded: April 8, 2013.

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

Borwein, J. (2008). Implications of experimental mathematics for the philosophy of mathematics. In B. Gold & R. Simon (ed). *Proof and other dilemmas:*

Mathematics and philosophy. (pp. 33-60). Mathematical Association of America: Washington, D.C.

Britt, M.A. & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction*, 20, 485-522.

Bromme, R., Kienheus, D., & Porsch, T. (2009). Who knows what and who can we believe? Epistemological beliefs are beliefs about knowledge (mostly) obtained from others. In L. Bendixen & F. Feucht (Eds.) *Personal epistemology in the classroom: Theory, research, and implications for practice*. (pp. 163-193). Cambridge: Cambridge University Press.

Buldt, B., Loewe, B., & Muller, T. (2008). Towards a new epistemology of mathematics. *Erkenntis*, 68, 309-329.

CadwalladerOlsker, T. (2011). What do we mean by mathematical proof? *Journal of Humanistic Mathematics*, 1, 33-60.

Cellucci, C. (2009). Indiscrete variations on Gian-Carlo Rota's themes. In E. Damiani et al (eds.) *From combinatorics to philosophy*, (pp. 211-228). Springer.

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

Chi, M. & Van Lehn, K. (2012). Seeing deep structure from the interactions of surface features. *Educational Psychologist*, 47, 177-188.

Chinn, C., Buckland, L., & Samarapungavan, A. (2011). Expanding the dimensions of epistemic cognition: Arguments from philosophy and psychology. *Educational Psychologist*, 46, 141-167.

Coady, C. (1992). *Testimony: A philosophical study*. Oxford: Oxford University Press.

Coe, R. & Ruthven, K. (1994). Proof practices and constructs of advanced mathematics students. *British Educational Research Journal*, 20, 41-53.

Common Core Content Standards Initiative (2012). Common Core State Standards initiative: Standards for mathematical practice. Downloaded from: http://www.corestandards.org/Math/Practice. Last downloaded: October 27, 2013.

Echeverria, J. (1996). Empirical methods in mathematics. A case study: Goldbach's Conjecture. In G. Munevar (Ed.) *Spanish Studies in the Philosophy of Science*. (pp, 19-56). Dordrecht: Kluwer.

Davis, P. (1972). Fidelity in mathematical discourse: Is one and one really two? *American Mathematical Monthly*, *79*, 252-263.

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

De Millo, R. A., Lipton, R. J., & Perlis, A. J. (1979). Social processes and proofs of theorems and programs. *Communications of the ACM, 22*, 271-280.

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

de Villiers, M. (2004). The role and function of quasi-empirical methods in mathematics. *Canadian Journal of Science, Mathematics, and Technology Education*, 4, 397-418.

Devlin, K. (2003). The shame of it. From *Devlin's Angle*, an on-line publication of the Mathematical Association of America. Downloaded from: http://www.maa.org/devlin/devlin_archives.html. Last downloaded: April 7, 2013.

Dreyfus, T. (1990). Advanced mathematical thinking. In P. Nesher & J. Kilpatrick (Eds.) *Mathematics and cognition: A research synthesis by the International Group for the Psychology of Mathematics Education*. Great Britain: Cambridge University Press.

du Sautoy, M. (2004). The Music of the Primes: Why an unsolved problem in mathematics matters. London: Harper Perennial.

Fallis, D. (1997). The epistemic status of probabilistic proofs. *Journal of Philosophy*, 94, 165-186.

Fallis, D. (2003). Intentional gaps in mathematical proofs. *Synthese*, 134, 45-69.

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

Franciso, J. & Maher, C. (2005). Conditions for promoting reasoning in problem solving: Insights from a longitudinal study. *Journal of Mathematical Behavior,* 24, 361-372.

Garofalo, J. (1989). Beliefs and their influence on mathematical performance. *Mathematics Teacher, 82,*502–505.

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

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

Grcar, J. F. (2013). Errors and corrections in mathematics literature. *Notices of the American Mathematical Society, 60*, 418-425.

Greene, J., Torney-Purta, J., & Azevedo, R. (2010). Empirical evidence regarding relations among a model of epistemic and ontological cognition, academic performance, and educational level. *Journal of Educational Psychology*, 101, 234-255.

Hanna, G. (1991). Mathematical proof. In D. Tall (Ed.) *Advanced Mathematical Thinking.* Dordrecht: Kluwer.

Hanna, G. (1995). Challenges to the Importance of Proof. *For the Learning of Mathematics,* 15(3), 42-50.

Hanna, G. & Barbeau, E. (2008). Proofs as bearers of mathematical knowledge. *ZDM,* 40, 345-353.

Harel, G. (2001). The development of mathematical induction as a proof scheme: A model for DNR-based instruction. In S. Campbell & R. Zazkis (Eds.) *Learning and teaching number theory*. New Jersey: Ablex Publishing Company.

Harel, G. & Sowder, L. (1998). Students proof schemes. *Research in Collegiate Mathematics Education, 3*, 234-282.

Harel, G. & Rabin, J. (2010). Teaching practices associated with the authoritarian proof scheme. *Journal for Research in Mathematics Education*, 41, 14-19.

Harel, G. & Sowder, L. (2007). Towards a comprehensive perspective on proof. In F. Lester (Ed.), *Second handbook of research on mathematical teaching and learning*. Washington, DC: NCTM.

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

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

Herbst, P. & Balacheff, N. (2009). Proving and knowing in public: The nature of proof in a classroom. In M. Blanton, D. Stylianou, & E. Knuth (Eds.) Teaching and learning proof across the grades: A K-16 perspective (pp. 40-64). New York: Routledge.

Hodson, J. (1996). Laboratory work as scientific method: Three decades of confusion and distortion. *Journal of Curriculum Studies*, 28, 115-135.

Hofer, B. (2000). On dimensionality and disciplinary differences in personal epistemology. *Contemporary Educational Psychology*, 25, 378–405.

Hofer, B. & Pintrich, P. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67, 88–140.

Housman, D. & Porter, M. (2003). Proof schemes and learning strategies of above average students. *Educational Studies in Mathematics*, 53, 139-158.

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

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

Inglis, M. & Alcock, L. (2013). Skimming: A response to Weber and Mejia-Ramos. *Journal for Research in Mathematics Education*, 44, 471-474.

Inglis, M. & Mejia-Ramos, J.P. (2009). The effect of authority on the persuasiveness of mathematical arguments. *Cognition and Instruction, 27*, 25-50.

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

King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco: Jossey-Bass.

Knuth, E. (2002). Secondary school mathematics teachers' conceptions of proof. *Journal for Research in Mathematics Education*, 33(5), 379-405.

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

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

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

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

Lakatos, I. (1978). *Mathematics, science, and epistemology*. Cambridge, MA: Cambridge University Press.

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.

Lester, F. K., & Garofalo, J. (1987). The influence of a*ff*ects, beliefs, and metacognition on problem solving behavior: Some tentative speculations. Paper presented at The Annual Meeting of the American Educational Research Association, Washington, DC.

Lipton, P. (1988). The epistemology of testimony. *Studies in History and Philosophy of Science*, 29, 1-31.

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

Marriotti, M. (2006). Proof and proving in mathematics education. In A. Guttierez & P. Boero (Eds.) *Handbook of research in mathematics education: Past, present, and future [PME 1976-2006]*. (pp. 173-204). Rotterdam: Sense Publishers.

Martino, A. M., & Maher, C. A. (1999). Teacher questioning to promote justification and generalization in mathematics: What research practice has taught us. *The Journal of Mathematical Behavior, 18*(1), 53-79.

McKnight, C., Magid, M., Murphy, T.J., & McKnight, M. (2000). *Mathematics education research: A guide for the research mathematician*. American Mathematical Society: Washington, D.C.

Moore, R.C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics*, 27, 249-266.

Mejia-Ramos, J.P. & Weber, K. (in press). How and why mathematicians read proofs: Further evidence from a survey study. To appear in *Educational Studies in Mathematics*.

Muis, K. R. (2004). Personal epistemology and mathematics: A critical review and synthesis of research. *Review of Educational Research*, 74, 317-377.

Muis, K. (2008). Epistemic profiles and self-regulated learning: Examining relations in the context of mathematics problem solving. *Contemporary Educational Psychology*, 33, 177-208.

Muis, K. R., & Franco, G. M. (2009). Epistemic beliefs: Setting the standards for self-regulated learning. *Contemporary Educational Psychology, 34*, 306-318.

Müller-Hill, E. (2010). *Die epistemische Rolle formalisierbarer mathematischer Beweise. Formalisierbarkeitsbasierte Konzeptionen mathemaischen Wissens und mathematischen Rechtfertigung innerhalb einer sozioempirisch informierten Erkenntnistheorie der Mathematik.* Unpublished doctoral dissertation. Rheinische Friedrich-Wilhelms-Universitat Bonn.

National Council of Teachers of Mathematics (2000). *Principles and Standards for School Mathematics*. Reston, VA: NCTM.

Nathanson, M. (2008). Desperately seeking mathematical truth. *Notices of the American Mathematical Society, 55*(7), 773.

Nathanson, M. (2009). Desperately seeking mathematical proof. *Mathematical Intellegencer*, 31, 8-10.

Osborne, J., Collins, S., Ratcliffe, M., Millar, R., & Duschl, R. (2003). What "ideas-about-science" should be taught in school science? A Delphi study of the expert community. *Journal of Research in Science Teaching*, 40, 692-720.

Otte, M. (1994). Mathematical knowledge and the problem of proof. *Educational Studies in Mathematics*, 26, 299-321.

Paseau, A. (2011). Mathematical instrumentalism, Godel's theorem, and inductive evidence. *Studies in History and Philosophy of Science*, 42, 140-149.

Perry, W. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. San Francisco: Jossey-Bass.

Polya, G. (1954). *Mathematics and plausible reasoning: Induction and analogy in mathematics*. Princeton University Press: Princeton, NJ.

Putnam, H. (1979). *Philosophical papers Volume 1: Mathematics, matter, and method*. Cambridge: Cambridge University Press.

Rasmussen, C. & Marrongelle, K. (2006). Pedagogical content tools: Integrating student reasoning and mathematics into instruction. *Journal for Research in Mathematics Education*, *37*, 388-420.

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

Rav, Y. (2007). A critique of formalist-mechanist version of the justification of arguments in mathematicians' proof practices. *Philosophia Mathematica*, *12*, 291-320.

Recio, A.M. & Godino, J.D. (2001). Institutional and personal meanings of proof. *Educational Studies in Mathematics*, 48(1), 83-99.

Rehmeye, J. (2013). A theorem in limbo shows that QED is not the last word in a mathematical proof. *ScienceNews Website*, 25th March 2013, http://t.co/KSgikZ2Ioa. Last downloaded: April 8, 2013.

Rota, G-C. (1991). The concept of mathematical truth. *Review of Metaphysics*, 44, 483-494.

Schoenfeld, A.H. (1983). Beyond the purely cognitive: Belief systems, social cognitions, and metacognitions as driving forces in intellectual performance. *Cognitive*

*Science, 7,*329–363.

Schoenfeld, A. H. (1985). *Mathematical Problem Solving*. Orlando, FL: Academic Press.

Schoenfeld, A.H. (1992). Learning to think mathematically: Metacognition, problem solving, and sense making in mathematics. In D. Grouws (Ed.) *Handbook of research on mathematical thinking and learning*. New York: Macmillan.

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

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82*, 498-504.

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*, 4-36.

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

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

Shanahan, C., Shanahan, T., & Misischia, C. (2011). Analysis of expert readers in three disciplines: History, mathematics, and chemistry. *Journal of Literacy Research*, 43, 393-429.

Simpson, A. (1995). Focusing on student attitudes to proof. *Teaching and Learning Undergraduate Mathematics Newsletter*, 3.

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

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

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

Stylianides, A. & Stylianides, G. (2009). Proof construction and evaluation. *Educational Studies in Mathematics*, 72, 237-253.

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

Stylianou, D. A. (2002). On the interaction of visualization and analysis—The negotiation of a visual representation in problem solving. *Journal of Mathematical Behavior*, *21*, 303–317.

Tao, T. (2007). What is good mathematics. arXiv:math/0702396 [math.HO].

Thurston, W. P. (1994). On proof and progress in mathematics. *Bulletin of the American Mathematical Society, 30*(2), 161-177.

Tymoczko, T. (Ed.) (1986). *New directions in the philosophy of mathematics*. Secaucus, NJ: Birkhauser.

Vinner, S. (1997). The pseudo-conceptual and pseudo-analytic thought processes in mathematical learning. *Educational Studies in Mathematics*, 34, 97-129.

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

Weber, K. (2003). Students' difficulties with proof. In A. Selden and J. Selden (eds.) *Research Sampler,* 8. Available from the following MAA website on the teaching of learning of mathematics: http://www.maa.org/t_and_l/sampler/research_sampler.html. Last downloaded April 18, 2013.

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

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

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

Weber, K. (2013). On the sophistication of naïve empirical reasoning: Factors influencing mathematicians' persuasion ratings of empirical arguments. *Research in Mathematics Education*, 15, 100-114.

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

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

Weber, K. & Mejia-Ramos, J.P. (2013a). On influence of sources in the reading of mathematical text: A reply to Shanahan, Shanahan, and Misischia. *Journal of Literacy Research*, 45, 87-96.

Weber, K. & Mejia-Ramos, J.P. (2013b). On mathematicians' proof skimming: A reply to Inglis and Alcock. *Journal for Research in Mathematics Education*, 44, 464-471.

Wineberg, S. (1990). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, 83(1), 73-87.

Wu, H. (1997). The mathematics education reform: Why you should be concerned and what you can do. *American Mathematical Monthly*, 104, 946-954.

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

Yackel, E. & Hanna, G. (2003). Reasoning and proof. In J. Kilpatrick, W. Martin, & D. Schifter (Eds.) *A research companion to the NCTM Standards* (pp. 227-236). Washington, DC: NCTM.