There is a robust body of research demonstrating that when students are asked to justify a mathematical assertion, they will frequently generate empirical arguments to do so. They also sometimes claim a deductive argument does not supply them with certainty that the assertion is correct. Mathematics educators frequently attribute this to students having deficient standards of conviction. In this paper, we illustrate another theoretical account. Students might believe that they lack the cognitive capacity to produce a superior argument to an empirical argument or to verify that a deductive argument is correct.

Key words: Expectancy value; Justification; Proof

In this paper, we consider two robust findings from the research literature on proof. First, when students are asked to justify a mathematical assertion, they frequently do so by generating an empirical argument (e.g., Coe & Ruthven, 1994; Healy & Hoyles, 2000; Recio & Godino, 2001). Second, when some students are presented with a valid deductive argument (i.e., a proof) in support of a mathematical assertion, they will claim not to be certain that the claim is correct (e.g., Chazan, 1993; Fischbein, 1982; Morris, 2002). The research questions addressed in this proposal question why is this the case.

Theoretical Perspective

The Proving as Convincing Research Paradigm

According to Stylianides, Stylianides, and Weber (in press), much of the vast mathematics education on justification and proof uses the metaphor that proving is tantamount to convincing, in which a critical goal of the Proving as Convincing paradigm is to identify the types of arguments that students find convincing. A common methodology to do so is to provide students with justification tasks, classify the students’ justifications by the types of arguments that they use to support the assertion (e.g., empirical generalizations, logical deduction, appeals to authority), and then infer that students presented these arguments because those are the types of arguments that are most convincing to students. Consequently, when students justify claims by empirical arguments, these researchers infer that students gain certainty on behalf of empirical arguments, or at least that empirical arguments are students’ preferred mode of justification. Researchers find this inference undesirable on the grounds that mathematicians obtain conviction from deductive arguments or proofs, not from empirical arguments (c.f., Harel & Sowder, 2007). Another common methodology in the Proving as Convincing paradigm is to present students with a proof of an assertion and then see if they behave as if they still harbor some doubt about that assertion. When students do, researchers also interpret this as undesirable because mathematicians, unlike these students, gain certainty from proof (c.f., Harel & Sowder, 2007).

Recently, the methodologies described above, as well as the epistemological assumptions that underpin them, have been challenged by some researchers. We briefly review some of the challenges here. First, the notion that mathematicians always gain certainty in mathematical assertions on behalf of proofs and that mathematicians never gain certainty on behalf of empirical arguments has been shown to be an oversimplification that does not accord with
modern mathematical practice (e.g., Weber, Inglis, & Mejia-Ramos, 2014). Second, studies in which students are interviewed by their purportedly non-normative judgments, such as claiming to be convinced by an empirical argument or retain doubts in a statement that has been proven, have found students to appear more sensible and mathematically normative than the common interpretations for the literature suggest (e.g., Bieda & Lepak, 2014; Stylianides & Al-Mourani, 2010; Weber, 2010). Third, the results from some empirical studies have suggested that assertions that are difficult to prove questions were more likely to elicit empirical arguments because proofs might be too difficult for students to generate (e.g., Knuth, Choppin, & Bieda, 2009; Stylianides & Stylianides, 2009).

**Our Alternative Expectancy Value Account**

In the educational psychology literature, expectancy value theories are broad theories of motivation that study the relationship of beliefs, values, and goals with action (Eccles & Wigfield, 2002). The central goal of this paper is to show how three central constructs in expectancy theories—values, costs, and likelihood of success—can be used to qualitatively account for students’ willingness or refusal to seek proofs of conjectures that they form.

Before sketching out our theory, we highlight an important distinction made by Bandura (1997) on two different expectancy beliefs, a difference that until now generally has not been taken accounted for in the research literature on proof. Bandura distinguished between *outcome expectations*—beliefs that certain behaviors will lead to certain outcomes—and *efficacy expectations*—beliefs that whether one has the capacity to perform the behaviors to achieve these outcomes. When students produce an empirical argument, we argue that we need to distinguish between two accounts: (i) the student believes the empirical argument provides certainty and (ii) the student believes that he or she does not have the capacity to produce a better argument. (i) is an outcome expectation and (ii) is an efficacy expectation. As we noted earlier, (i) is a common inference in the literature and suggests that students are not aware of the limitations of empirical arguments. However, we will argue that (ii) is sometimes the more accurate account. Similarly, if students are not certain of a claim after reading a proof of the claim, two possible accounts are (i) the student does not think that proofs can provide certainty or (ii) the student questions whether he or she has the ability to be certain that a proof is correct. Again, researchers often infer (i), an outcome expectation, but we believe that (ii), an efficacy expectation, often provides a more accurate account.

**Value.** When students are asked to justify a mathematical assertion, we assume that students want to achieve a high level of conviction for why an assertion is true. Our premise is that the *value* that students put on obtaining high levels of conviction will influence how hard students will work to obtain a better justification and whether they will settle for a non-optimal explanation as good enough. These values can be externally or internally imposed. Externally, submitting a false conjecture for an assignment (or a true conjecture that is not adequately supported) can result in the student receiving a negative outcome, such as a low grade for a course, or otherwise result in some other consequence that would hurt their academic or career goals. Eccles and Wigfield (2002) referred to this as Utility Value. Internally, students may wish to resolve a genuine curiosity as to whether a conjecture is true or false; Harel (1998) referred to as Intellectual Need. Also, a student may wish to avoid presenting a false conjecture to his or her peers or teacher as this would harm his or her mathematical self-image; Eccles and Wigfield (2002) referred to as Attainment Value. The higher the perceived value of determining the truth of a conjecture, the more likely a student is to work to produce a justification that bestows a high level of conviction.
Cost. Once a student has an imperfect justification for why a mathematical statement is true, the time spent searching for a superior justification, including a proof, is time that could be spent on other enjoyable activities, such as talking to friends. Here it is important to observe that the cost of searching for a proof may depend heavily on the individual and the situation. For many students, attempting to write a proof is a painful activity that brings up feelings of intellectually inferiority (e.g., Weber, 2008). Other students and mathematicians may enjoy seeking a proof, so engaging in the activity of proving is an end in and of itself—having what Eccles and Wigfield (2002) referred to as a high Intrinsic Value. In the latter case, individuals will avoid having someone else show them a proof because it would deny them the opportunity to search for that proof themselves. The higher the cost, the less likely it is that a student will work to produce a proof.

Likelihood of success. In deciding whether to pursue an objective, students will make a subjective estimation of how likely they are to achieve that objective. If students have an imperfect justification for why a mathematical statement is true, they are unlikely to seek a proof unless they believe that there is a reasonable chance that their search will be successful. If students think it is highly unlikely that they can prove a statement, they may simply not seek one and settle for an empirical justification, even if they are aware that a proof, in principle, bestows more conviction than the empirical argument.

Earlier, we argued that when a student submitted empirical arguments, researchers have often drawn the conclusion of that student having a limited epistemology. The researcher made the inference that the student believed that empirical arguments can bestow certainty, or at least more conviction than deductive arguments. Here we offer three other possibilities: (i) students might not be interested in being certain of the conjecture in question, (ii) they might find searching for a proof to be an unpleasant endeavor that is not worth the effort, or (iii) they might settle for the empirical argument because they believe that they lack the capacity to find a proof.

Methods

Rationale

In this study, we asked prospective and inservice secondary mathematics teachers to work on challenging problems that invited students to use empirical reasoning to make conjectures. We asked participants to share with us both their answer to the problem and a justification for why they believed their answer was correct. Following Stylianides and Stylianides (2009), we also asked participants how confident they were in their answers on a scale of 0 through 100. In cases where participants gave a response of less than 100, we asked them why they retained doubt about their answer, what further evidence could give them more confidence that their solution was correct, and why they were not seeking that evidence.

The goal of this study was twofold. The first is to argue that the Proving as Convincing research paradigm often does not offer an accurate account of why students will justify their conjectures with empirical arguments. Consistent with the research literature, in our study, we found that participants frequently justified their answers with empirical arguments. If participants were doing so because they held undesirable standards of conviction, we would expect the participants to have certainty, or at least a high degree of confidence, in their answers on behalf of these empirical arguments. At a minimum, the participants should aver that the empirical arguments provided them with at least as much confidence as a proof could provide. However, this is not what we observed. Second, we want to illustrate how our expectancy value model can explain why participants offered the empirical arguments that they did. Taking into account the
participants’ perceived value of obtaining complete conviction, cost anticipated in searching for a proof, and likelihood of success in finding a proof allows us to provide coherent accounts of why students produced and settled for empirical arguments. Indeed, we will further argue that in many cases when participants offer empirical arguments, participants are neither behaving irrationally nor inconsistent with mathematical practice.

Research Context

This study occurred at a large state university in the northeast United States. The data was collected in the first six weeks of a content-based course for prospective and practicing teachers of secondary mathematics focusing on problem solving. The aim of the course was for these teachers to develop the mathematical knowledge and dispositions to solve mathematical problems effectively and to enable these teachers to help their future students to do so. The course met weekly and was co-taught by the first two authors of this paper.

For the first six weeks of the course, class meetings were comprised of the students solving challenging problems in groups for about one hour, presenting their solutions to the larger class, and then reflecting on how this experience related to the nature of mathematical problem solving and the teaching of mathematics. The data from the paper was collected in four of these six meetings.

Our analysis focuses on the 11 prospective or practicing mathematics teachers who completed the course, all of whom claimed to have completed some courses in advanced mathematics. The students worked in four groups, which we call Group A, B, C, and D, which were mostly stable throughout the semester.

Materials. The data was collected when the students worked on four problems, two of which have multiple parts. These problems were chosen such that the answer to the questions could be conjectured via empirical reasoning but each answer could also be justified by deductive argumentation or proof. Problems 1, 2, 3, and 4 were given in the first, second, fourth, and sixth week of the course, respectively. (The problems in week 3 and week 5 were a geometry and a modeling problem respectively. These were not included in our analysis because they did not permit empirical generalizations).

Procedure. For the first six weeks of the semester, class began by having students work collaboratively on the mathematical problems. They were given one hour to work on the problems. If a group had agreed upon an answer to the problem, they could raise their hand to discuss the answer with one of the two course instructors. When this occurred, the instructors would interview the students about their solutions. Otherwise the instructors would circulate the room observing the students’ behavior and answering questions that the students may have had about the meaning or interpretation of the problem. The instructors would not, however, provide hints or assistance to the student or confirm that their answers were correct or that the students were on the right track. After 60 minutes had elapsed, if a group did not discuss their solution with the instructors (which was usually the case), the instructors would then go to each group and ask them to discuss the answer that they obtained.

The interviews between each group of students and a course instructor were semi-structured and audiotaped. The instructor photographed the students’ written solution. The instructor first asked the student what their solution was and then asked the students to explain and justify their solution. At this stage, the instructor sometimes asked for clarification if he or she was not sure that she understood the students’ arguments. The interviewer then asked the students to state how certain their solution was correct on a scale of 0 through 100. If any students gave a response of less than 100, they were asked why they were not certain in their solution, what further evidence
could give them more confidence in their answer, and why they were not seeking this evidence. In these discussions, the interviewer carefully distinguished between students’ answers and their justifications. The interviewer would explicitly say that he or she wanted the students’ confidence about the correctness of the answer, not the correctness or permissibility of the justification. Interviews typically lasted between five and ten minutes.

Problem 3 had two sub-questions and Problem 4 had three sub-questions. For these class meetings, each sub-question was treated as a separate problem. (i.e., the instructor would go through the protocol for each explanation that each group provided for one of the problems). Hence, each group was given the opportunity to supply seven answers—one for Problem 1, one for Problem 2, two for Problem 3, and three for Problem 4. For some problems, the groups disagreed on the answer and offered multiple justifications. In these cases, each answer was treated separately. In total, 31 answers were offered.

**Analysis.** All of the audio-recorded data was transcribed. The analysis proceeded through three stages. In the first stage of the analysis, we categorized each justification as empirical or deductive. For the sake of brevity, we do not discuss how we assigned these codes here. In the second stage, we determined the confidence that individual students had in their group’s answer. In the third stage, we engaged in thematic analysis (Braun & Clarke, 2006) to categorize the reasons that (1) students expressed certainty in an answer, (2) why students expressed some doubt in their answer (i.e., gave a conviction score of less than 100), and (3) why they did not seek more evidence when they gave a conviction score of less than 100.

**Results**

Table 1 presents a summary of the arguments that students’ produced as well as how confident that the students were that the argument was correct. Table 1 reveals that the majority (18 of 31) of students’ justifications for their answers were empirical in nature, replicating the result from the research literature that students frequently justify by empirical arguments. However, when students justified their answer with an empirical argument, they usually retained some doubt about the correctness of their answer, giving a confidence rating of less than 100 in 32 of their 39 (82%) of their ratings. Consequently, these data indicate that students’ propensity to justify their mathematical claims with empirical evidence does not imply they believe that empirical evidence can provide certainty that a mathematical claim was true.

<table>
<thead>
<tr>
<th></th>
<th>Deductive Justification</th>
<th>Empirical Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of justifications</td>
<td>13</td>
<td>18</td>
</tr>
<tr>
<td>Students with a confidence level of 100 for the justification</td>
<td>27</td>
<td>7</td>
</tr>
<tr>
<td>Students with a confidence level of less than 100</td>
<td>6</td>
<td>32</td>
</tr>
<tr>
<td>Average confidence level</td>
<td>99.4</td>
<td>68.9</td>
</tr>
</tbody>
</table>

Table 1. Confidence level of students in deductive and empirical justifications.

To understand why students gave the responses that they did, consider Group B’s solution and justification for Problem 1, where Billy offered the incorrect answer of \( \frac{3}{4} \).

**Problem 1:** Suppose there are a row of squares with each square numbered 1, 2, 3, and so on. You start on square 1. You flip a coin. If you get heads, you move up one space. If you get tails, you move up 2 spaces. You repeat this process indefinitely. What is the probability that, at some point, you are on square 25?

[1] Int1: Okay. On a scale of 0 through 100, how confident are you that your solution is correct?
That the answer is \( \frac{3}{4} \).

Bob: \( \frac{3}{4} \). [Laughter]

Int2: 75. Is that a fair number?

Bob: Wasn’t it like the same pattern when he worked it out with 3? So first he worked it out with like 3 and then he tried from like 23 to 25 and then he got like the pattern was still continuing. So that kind of convinced me too.

Brenda: I’d say 98.

Bob: I’m confident. So I’d say 98 too.

Int2: So none of you said 100. So why didn’t you say 100? What doubts do you still have?

Billy: For me, I’m a little doubtful because I know like the long way of finding the answer of finding the exact, like the number of heads, the number of tails, the number of choices, blah, blah, blah, but that’s way too long. So I’m trying to shorten my way and find the answer. Which means I’m like eyeballing it to say. Because of that, I’m not 100% sure but I’m pretty confident that this is the right way.

Int2: So are you saying you are willing to take a slight risk that this might not be right to avoid the longer process… I’m sorry do you have anything that you would like to add to that Brenda?

Brenda: No, I mean, I’m pretty confident. I would say, because I don’t know, if this was a question, I would probably stop at that and be like, I’m going with this question. If it was like an exam and it was pass or fail, this would be my final answer.

Int2: Would there be any evidence or further work that would make you more confident?

Brenda: I guess like what Billy was saying, maybe doing the whole long way, going out but realistically, I don’t know. I’m kind of lazy in that aspect.

Int2: You started to answer this, but why wouldn’t you seek that evidence…

Billy: It’s too long.

Brenda: Yeah.

In this excerpt, we again see the students claiming that they did not find the empirical argument to be fully convincing [lines 6-8]. Billy and Brenda said they would be more convinced by a deductive argument [9, 13] yet both said that they would not seek the deductive argument because it would not be worth the effort [15-16]. What is striking is that Billy claimed that he knew how to produce a proof [9] and he felt that the likelihood of obtaining a proof was high [9]. Even though Billy’s confidence in the answer that he offered was relatively low [3], he did not produce this proof. For Brenda and Billy, the value of raising his confidence level was not worth the cost of seeking a lengthy proof.

Group B was told that their answer was incorrect and they were asked to continue working on the problem. Later, Group B obtained the correct answer, justifying it with a deductive argument (the one surveyed by Billy in [9]).

Int: Let me ask you the same question. How confident are you on a scale of 0 through 100 that that is correct?

Billy: [pause] I’ll go with 100 for this one.

Brenda: 100.

Bob: I’ll go with 99.

Int: 100. Why are you certain?

Billy: For no reason. [laughs] Because I worked, I calculated, I’ve come up with a formula then I applied it to getting to 3 from 1 and getting to 5 from 1 and it worked out.

Int: So you had the logic for your formula…

Billy: Yeah.

Int: And you checked it for small cases.

Billy: Yes.

Int: Bob, is there a reason that you still have a scintilla of doubt?

Bob: I don’t know. [pause] It’s just I’m kind of shaky about the whole choose thing. But the way he kind of presented it, I kind of like, did it in an old-fashioned kind of list and I got the same maximum of...
like 24 and minimum of like 12 flips. So like I’m agreeing with what he’s saying. What’s kind of stopping me is the whole choose thing and then we’re assuming that order doesn’t matter, right?

We observe two things in this transcript. First, Billy claims certainty in his answer [18], but this is not just because he verified the logic in his argument. It is also because he checked that his argument worked in the case of a simple example [22]. This was a common occurrence in this study. Second, Bob obtains a high degree of confidence but not certainty in Group B’s answer [20] even though he can follow Billy’s argument [28]. Our interpretation of the explanation in [28] is that Bob felt that he lacked the background in combinatorics (“I’m kind of shaky about the whole choose thing”) to ensure that there was not a mistake in Billy’s reasoning.

In the full paper, we will illustrate that the themes discussed here—that students settled for empirical arguments due to a lack of motivation, that they obtained certainty in deductive arguments by coordinating them with empirical evidence, and that students who expressed doubt in a proven statement did so because they were not confident in their ability to verify a proof, were common occurrences in this study.

Discussion

In our study, we found that, consistent with the research literature, students often justified their mathematical assertions with empirical arguments and sometimes expressed doubt in a statement after reading (and accepting) a proof of that statement. The Proving as Convincing perspective in the mathematics education literature (as described by Stylianides, Styliandies, & Weber, in press) can account for some of these occurrences. There were instances in which a student genuinely appeared to gain certainty from an empirical argument. However, for most instances, the participants submitted empirical arguments because they lacked the motivation to seek a proof or doubted that they could produce a proof if they sought one. Consequently, we propose that our broader Expectancy Value framework can offer a more accurate account for why students offer the justifications that they do. One implication for mathematics education researchers is that they should ask students to express the level of confidence that they have in their solution after they supply a justification (e.g., Stylianides & Stylianides, 2009; Weber & Mejia-Ramos, 2015).

As a final caveat, we note that our study occurred with mathematics teachers who had experience in proof-based mathematics courses. This population appears to differ from other populations in how they construct and evaluate justifications (e.g., Iannone & Inglis, 2010; Weber, 2010). Hence, we do not claim at this stage that our empirical findings will generalize to these populations, only that the theoretical account that we provide in this paper should be taken into account when conducting research with this population.
References