Students' Examining and Making Sense of Proofs by Mathematical Induction

Samuel D. Reed
Middle Tennessee State University

Jordan Kirby
Middle Tennessee State University

Sarah K. Bleiler-Baxter
Middle Tennessee State University

Students have demonstrated difficulties in adopting the various techniques of proof in their upper-level mathematics coursework. One of these techniques of proof which students struggle with is mathematical induction. In this paper, we present an analysis of three groups of students in an Introduction to Proofs course as they made sense of and analyzed two sample induction arguments. To aid in our analysis, we utilized Stylianides' (2007) definition of proof with three components (accepted statements, modes of argumentation, modes of argument representation) to describe the components of proof students focus upon when they encounter induction arguments for the first time in an Introduction to Proof course.

Keywords: Mathematical Induction, Intro to Proof, Proof, Proving

While the transition to formal proof is difficult for many (Moore, 1994; Harel & Sowder, 2007; Stylianides et al., 2017), students have shown that there are unique challenges to adopting individual techniques of proof (e.g., direct proof, contradiction). One of these techniques of proof which can be particularly troublesome for students is mathematical induction. There are a few potential reasons for students' trouble. First, induction relies on the *well-ordering principle*, which has puzzled mathematicians into the 20th century (Hellman, 2010). Second, and potentially more importantly to novices of proof writing, induction uses the conditional statement in unique ways relative to other techniques of proof (See Table 1), with an embedded conditional as part of the proof structure (i.e., the induction step) and with the notion that proof by induction is utilized when proving an infinite sequence of conditional statements. This, combined with induction's necessary reliance on a particular example (i.e., the base case), has the potential to leave students confused about both the form and function (Stylianides et al., 2007; 2016) of proofs by mathematical induction.

Table 1
Conditional Use Per Proof Technique

Proof Technique	Conditional Use
Direct Proof	$P \rightarrow Q$
Proof by Contraposition	$\sim Q \rightarrow \sim P$
Proof by Contradiction	$(P \text{ and } \sim Q) \rightarrow F$
Mathematical Induction	$[P(1)]$ and $[P(n) \rightarrow P(n+1)]$
Disproof by Counterexample	$T \rightarrow F$

Studies on students' use and understanding of mathematical induction are limited. Students have expressed that proofs by mathematical induction lack explanatory power (Stylianides et al., 2007), although Stylianides and colleagues (2016) found that students' example use promoted such understanding when proving conjectures with mathematical induction. In regards to the components of mathematical induction, students have demonstrated a lack of understanding about the necessity of the base case (Dubinsky, 1986; 1990; Stylianides et al., 2007), and use circular reasoning when describing the inductive hypothesis (Harel, 2001; Palla et al., 2012). Stylianides and colleagues (2007) surveyed over 100 preservice elementary and secondary mathematics teachers and analyzed their response to two sample induction arguments as to whether or not the proofs were mathematically valid. In their conclusion, the authors suggest that the two tasks that they used in their survey could also potentially be used within classroom contexts and that future research would benefit from investigating such in-class experiences with the tasks. We continue their line of research by investigating small-groups of students in an Introduction to Proof course as they work together to evaluate two sample induction arguments. Our goal is to understand, within a natural classroom setting, how students make decisions about mathematical induction. In particular, we seek answers to the following research question: When students are asked to evaluate sample proofs by mathematical induction, what do they base their decisions upon with respect to whether the induction argument is valid?

Theoretical & Analytic Framework

In order to frame the ways in which students evaluate and make sense of induction proofs, we utilized Stylianides' (2007) description of mathematical proof, which is parsed into three categories: (i) Set of Accepted Statements; (ii) Modes of Argumentation; and (iii) Modes of Argument Representation (p. 292). Our adaptation is summarized in Table 2 below.

Table 2

Proof Component Framework and Description (Stylianides, 2007, p. 292)

	Stylianides Definition	Example Quotes: Group Proving
Set of Accepted Statements (S)	● Definitions, axioms, theorems ● Prior proven results	Billy: "I think it's important to look at the base case. So we should check to make sure it works for N equals one."
Modes of Argumentation (A)	● Modus ponens/ Modus tollens ● Listing of all possible cases (if finite), or methods like contradiction, contrapositive, induction ● Logical derivations of statements	Zoren: "I'm saying you move to K plus one, the left side is multiplied by something, K plus one has to be greater than five. I mean, that's the left side, the right side. You're just multiplying it by two."
Modes of Argument Representation (R)	● Language to describe arguments ● Physical representation of arguments (i.e. pictorial or symbolic representation)	Billy: "It looks like a good proof. It does look like it works out due to the algebra. But they didn't check the very starting case. "

We used this framework as a way to categorize the statements made by students as they worked in small groups to evaluate sample induction arguments. Our coding allowed us to determine if students were basing their decisions upon accepted statements, modes of argumentation, and/or modes of argument representation. Through categorizing student responses in this way, we hoped to better understand what students find most relevant in their process of evaluating sample induction arguments.

Methodology

Task & Participants

The data for this study came from the 1st Author's (in progress) dissertation data on undergraduate students in an Introduction to Proofs course as they completed small-group proving tasks. The study took place at a large, public university, in the southeast United States. Of the 11 students in the study, 2 were women and the remaining 9 were men, and all students were either majoring or minoring in mathematics. The small-group proving task used in this study was modified from Stylianides et al. (2007) study on preservice teachers' understanding of proofs by mathematical induction. The task (replicated in Figure 1) prompted students to discuss two sample induction arguments and to explain why or why not the given arguments prove the stated claims. The task differs from Stylianides et al. (2007) in that the students were not asked to choose from a list of predetermined options on the proof's validity.

Figure 1

Sample Induction Arguments

For every $n \in \mathbb{N}$ the following is true: $1+3+5+\ldots+(2n-1)=n^2+3$

Proof:

First, I assume that the result is true for n=k.

This means that, $1+3+5+\ldots+(2k-1)=k^2+3$.

I check whether the result is true for n=k+1:

$1+3+5+\ldots+(2k-1)+(2k+1)=(k^2+3)+2k+1=(k^2+2k+1)+3=(k+1)^2+3$

True.

Therefore, $1+3+5+\ldots+(2n-1)=n^2+3$ for all $n \in \mathbb{N}$ by mathematical induction.

For every natural number $n \geq 5$ the following result is true:

$$1 \bullet 2 \bullet 3 \bullet \ldots \bullet (n-1) \bullet n > 2^n$$

Proof:

I check whether the result is true for $n=5$.

$$1 \bullet 2 \bullet 3 \bullet 4 \bullet 5 = 120 > 2^5 = 32$$

So, I assume the result is true for n=k : $1 \bullet 2 \bullet 3 \bullet \ldots \bullet (k-1) \bullet k > 2^k$

Now, I check whether the result is true for n=k+1.

$$\begin{aligned} 1 \bullet 2 \bullet 3 \bullet \ldots \bullet (k-1) \bullet k \bullet (k+1) &> 2^k \bullet (k+1) \\ &> 2^k \bullet 2 \qquad (\textit{because } k+1>2) \\ &= 2^{k+1} \end{aligned}$$

Therefore, $1 \bullet 2 \bullet 3 \bullet \ldots \bullet (n-1) \bullet n > 2^n$ for all $n \geq 5$.

Students in this course had already been introduced to direct and indirect modes of proof, but were yet to be introduced to mathematical induction. Before class, students were asked to examine the two sample arguments and to be ready to discuss their initial thoughts. Prior to launching this task, the instructor (3rd Author) introduced students to the definition of mathematical induction shown in Figure 2.

Figure 2

Definition of Mathematical Induction Presented to Students

- To prove an infinite sequence of statements $P(1), P(2), P(3), \ldots$ [that is, $P(n)$ for all $n \in \mathbb{N}$], we will need to prove the following:
- **Base case:** Prove $P(1)$ is true.
- **Induction step:** Prove $P(k) \Rightarrow P(k+1)$ for $k \in \mathbb{N}$.
- Then $P(n)$ is true for all $n \in \mathbb{N}$. Note: The statement $P(k)$ in the induction step is called the **induction hypothesis.**

Students then connected this definition to the domino analogy and graphic of S(1) → S(2) → … S(K) → S(K+1) →.... Thus, we were able to analyze students' initial conceptions and understandings of mathematical induction, whether they recognized the components of mathematical induction and could connect them to the definition, and whether they found this argument form convincing.

Data Analysis

Three videos, 19 minutes each, were transcribed by talk turn where small groups of students (3-4 students each) analyzed the induction arguments outlined in Figure 1. The three authors then independently applied Stylianides' (2007) framework (described above) to code each talk turn, where applicable. Each line could be coded with any combination of the three modes described in Stylianides' frame, and we decided to code a line if students appeared to be making a comment, asking a question, or making a decision related to the set of accepted statements (S), the modes of argumentation (A) or the modes of argument representation (R). After independent coding, the three authors met to come to unanimous agreements on each talk turn's code.

Limitations

This introduction to proof course was conducted entirely online due to Covid-19. As such, students worked in groups through their personal computers in Zoom breakout groups. We feel that this may have had an impact on their discussion when analyzing the sample arguments, as the difficulties of online learning and writing of proof are unknown. Further, no students were interviewed regarding their discussions and understanding of induction. We may have interpreted some student interactions in a manner different than they intended.

Results

We present the results for this study in two sections. First, we show how the various groups responded to Argument 1, in particular highlighting how they made sense of the base case (or lack thereof). Next, we present the various groups' responses to Argument 2, in particular highlighting how the groups of students made sense of the inductive step. Students in this study discussed the base case (for argument 1) in two main ways: (a) as something that needed to be confirmed as part of a checklist for performing mathematical induction, and (b) as a specific example that can aid in determining the truth value of the given claim. Likewise, students in this study focused on certain aspects of the inductive step (in argument 2): (a) algebra and the inductive hypothesis, and (b) why they switch variables between "N" and "K".

Base Case Theme 1: Confirming the Base Case: A Checklist

The first sample argument which students were asked to analyze is one for which the claim is not true, and the proof is also missing the base case. A student (or group of students) in all three groups recognized both of these inaccuracies in the proof and conjecture. Table 3 shows the percentage of the talk turns related to the base case in Argument 1 that were coded as S, A, or R.

Table 3

Percentage of talk turns related to the base case in Argument 1 coded as S, A, or R

	S	A	R	# Talk Turns in Arg 1 Related to Base Case
Group 1	43	57	0	7
Group 2	36	93	0	14
Group 3	39	77	31	13

In our analysis, we found that students' discussions around the first sample argument relied heavily on sets of accepted statements (S), seeing the base case as a necessary step to check from the given definition of induction as opposed to the first true statement in an infinite series of implications (i.e., S(1) → … → S(K) → S(K+1) → …). Example quotes from the various groups with this checklist-type language are shown in Table 4.

Table 4

Student Checklist-Type Language Quotes

	Group 1 - Zoran	Group 1- Eliza	Group 2 - Taylor	Group 3 - Billy
Student Quote	For induction proof, you have to start with proving like the N equals one case and then the N equals one case is false. So you can't even go on from there.	I think the first one is false because they didn't check for a base case.	Well, I don't know.... Aren't you supposed to prove the first step is true and then prove K. And then..	I think it's important to look at the base case. So we should check to make sure it works for N equals one

Base Case Theme 2: Disproving by Counterexample

While discussing the base case in Argument 1 some students also noted that the statement is not true in general. These discussions are shown in Table 5. Here their talk was different than when discussing the lack of a stated base case, as their attention turned to the validity of the claim given an initial value (e.g., n=1; n=4). In these instances, students argued that the given argument could not be a valid proof because the claim itself could be shown to be false through a single counterexample. Note, this rationale for why the proof is invalid does not require any understanding of proof by mathematical induction; however, the need that students felt above to check the base case within a proof by induction allowed students to quickly discover that the given claim was indeed false for that case (and hence in general).

Table 5

Students Disproving by Counterexample

	Group 1 - Zoran	Group 2 - Alex	Group 3- Billy and Tim
Student Quote	So I looked up the first one and it's not even true when you plug in values...Like I mean, if you plug in n=4, it will be 1+3+5+7, and that 1+3+5+7 is equal to 16, but if you look at the right side, n^2+3 is equal to 19. So without even thinking if it counts as induction, the statement is not even true. It's not true for n=1 either.	Yeah, and I don't really think there's a way to modify it to make it a proof because the statement is not true.	Tim: Does it work at [one] though? Billy: No, because 1would have to equal $1^2 + 3$. Tim: Yeah, so is the whole thing just false then? Billy: Yes.

Inductive Step Theme 1: Algebra and the Inductive Hypothesis

The second proof, a valid proof by mathematical induction, caused students to have some unique trouble. In particular, two of the groups had a detailed discussion about their struggle to understand the logical steps of the inductive hypothesis, similar to the results of Stylianides et al., (2007). Table 6 shows the percentage of the talk turns related to the inductive step in Argument 2 that were coded as S, A, or R.

Table 6

Percentage of talk turns related to the inductive step in Argument 2 coded as S,. A, or R

S	A	R	# Talk Turns

Group 1	50	75	16.67	12
Group 2	85.71	71.43	14.23	7
Group 3	25	100	100	4

In this table we see the heavy reliance on both "S" and "A" for groups 1 and 2. In both of these groups, there seemed to be a tension between the logical and algebraic steps (A) and whether the argument satisfied the criteria for induction (S) with respect to the inductive step. For example, consider the following exchange between members of Group 1 as they worked together to determine if the inductive step was satisfied:

> **Eliza**: I think they want you to assume that you're going to be using some additional power of two. I think is the assumption that they're making here and why it's OK for them to say that K+1 is greater than two, that you can replace K+1 with two, because the assumption is on the right hand side, we're operating with powers of two.
> **Zoren**: I mean, say, like the left hand side is like four and the right hand side is three. If you replace that three with something smaller, it doesn't change whether the equality is true or false? I guess that is the idea.
> **Eliza**: Yeah.. That's I think that's part of it.
> **Zoren**: My question was like, so what if it's like three? Three is greater than four? And then you replace four with like two or something smaller. That would definitely change it from false to true. I guess. So, like switching something on the right hand side does. It could change the truth value.

In this exchange we see the group trying to make sense of whether or not these algebraic manipulations are logically valid (i.e., focusing on "A"). To these students, there seems to be a tension between the logical and algebraic steps (A) and whether this satisfies the accepted criteria for induction and the inductive step (S). To Zoren, it is not clear that each algebraic manipulation is its own implication statement. We observed this issue with other students in the algebraic manipulations and assumptions of the inductive step.

Inductive Step Theme 2: Why Switch Variables?

Table 6 demonstrates that when Group 3 was discussing the inductive step, they were primarily focused upon "A" and "R." The following conversation depicts the team's dilemma, considering why a switch from "N" to "K" is useful and/or warranted for a valid argument by mathematical induction:

> **Joe:** Oh, these two. OK. Over here. I still don't like that, what do you call it, let's say the problem is written in N. And then the first thing you do is put it a K, then K is equal to N, I never really understood that like. Why why would you substitute, like wouldn't working with N be just
> **Billy**: So what's the question.
> **Joe**: I'm just trying to figure out. So. I guess they would. What is it, they had N equals K?
> **Billy**: Yeah
> **Joe**: But I guess that's just like a framing thing, so you can get N=K+1
> **Billy**: Yeah. I mean, I guess it's just so you separate the original argument which had N's with the ones you're modifying.

In this exchange, we see a tension between the mode of argumentation (A) and mode of argument representation (R) in the necessity of replacing N with K. We see that the student, Joe, is focusing on the logic of the substitution (A), while Billy points out that this is a useful way to make a distinction between the statement of the conjecture and the representation of the inductive step (R).

Discussion & Conclusion

We pursued this research in an effort to understand, within a natural classroom setting, how students make decisions about mathematical induction. By studying small-groups of students as they worked together to evaluate sample induction arguments, we were able to gain insight about the students' decision-making and sense-making in real time, and through peer-to-peer discourse. Methodologically, we found this approach quite revealing in terms of the understandings exposed through peer discussions.

When analyzing students' small-group evaluations of Argument 1, we were able to identify two ways that students discussed the base case: (a) as something that needed to be confirmed as part of a checklist for performing mathematical induction, and (b) as a specific example that can aid in determining the truth value of the given claim. Our coding scheme brought these two themes to light, as (a) students focused on their perception of the two-part definition of induction (i.e., check off the base case), which we captured through our S (set of accepted statements) codes, and (b) students utilized specific numeric examples, including n=1 from the base case, in order to justify why the given conjecture must be false, which we captured through our A (modes of argumentation) codes.

When analyzing students' small-group evaluations of Argument 2, we identified two themes in terms of the students' focus on the inductive step: (a) algebra and the inductive hypothesis, and (b) why they switch variables between "N" and "K". Our coding scheme was likewise useful in supporting the identification of these themes as (a) students wrestled with how the algebraic and logical underpinnings of an argument (e.g., debating the transitivity of inequalities in Group 1), which was captured with A codes, aligned with the accepted process of the inductive step (i.e., P(k) implies P(k+1)), which was captured with S codes, and (b) students pondered the utility of substituting "K" for "N" in terms of the benefit to the argument (A codes) and in terms of the benefit to the communication and representation of the argument (R codes).

Note that in Stylianides et al. (2007), the authors posit that "by breaking the boundaries of students' normal experience, the[se] two tasks set up situations where procedural knowledge of the induction method is not enough for successful performance" (p. 164). Our findings suggest through discussing these tasks with their peers, students persisted in relying on their understood ritual (procedure) of induction. However, these tasks also allowed students to make connections beyond their initial procedural understanding of the definition of induction (S codes) and to the important aspects of the argument (A) and the representation of the argument (R). We believe that using such tasks in classroom discourse, as depicted in this study, could be beneficial as a starting point for introducing mathematical induction. We suggest that future research investigate the possible use of these tasks in classroom settings without first introducing the formal definition of mathematical induction. In this way, students may be less prone to rely upon the accepted ritual, and the environment could potentially further students' non-procedural understanding of mathematical induction.

References

Dubinsky, E. (1986). Teaching mathematical induction: I.The Journal of Mathematical Behavior, 5, 305-317.

Dubinsky, E. (1990). Teaching mathematical induction: II. The Journal of Mathematical Behavior, 8, 285-304.

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

Harel, G., & Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. *Second handbook of research on mathematics teaching and learning*, *2*, 805-842.

Hellman, H. (2010). *Great feuds in mathematics: Ten of the liveliest disputes ever*. John Wiley & Sons, Inc.

Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics*, *27*(3), 249-266.

Palla, M., Potari, D., & Spyrou, P. (2012). Secondary school students' understanding of mathematical induction: Structural characteristics and the process of proof construction. *International Journal of Science and Mathematics Education*, *10*(5), 1023-1045.

Stylianides, G. J., Sandefur, J., & Watson, A. (2016). Conditions for proving by mathematical induction to be explanatory. *The Journal of Mathematical Behavior*, *43*, 20-34.

Stylianides, A. J. (2007). Proof and proving in school mathematics. Journal for Research in Mathematics Education, 38(3), 289–321.

Stylianides, G. J., Stylianides, A. J., & Philippou, G. N. (2007). Preservice teachers' knowledge of proof by mathematical induction. *Journal of Mathematics Teacher Education*, *10*(3), 145-166.

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